XXI Workshop de Investigadores en Ciencias de la Computación

25 y 26 de abril de 2019 – San Juan – Argentina

LIBRO DE ACTAS

Facultad de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de San Juan
Red de Universidades con Carreras de Informática (RedUNCI)

XXI Workshop de Investigadores en Ciencias de la Computación -WICC 2019: libro de actas / Compilado por Nelson Rodríguez, María Murazzo, Manuel Ortega, María I. Lund. - 1a ed . - San Juan: Editorial UNSJ, 2019.

CD-ROM, PDF
Archivo Digital: descarga y online

ISBN 978-987-3984-85-3

1. Informática. 2. Informática Educativa. 3. Innovación Tecnológica. I. Rodríguez, Nelson, comp.
CDD 005 - 1118 páginas

ISBN 978-987-3984-85-3

9 789873 984853

Universidad Nacional de San Juan (UNSJ)

Rector
Oscar Nasisi

Decano de FCEFyN
Rodolfo Bloch

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

Coordinador Titular
Pesado Patricia (UNLP)
Coordinador Alterno
Estayno Marcelo (UNLZ)

Coordinadores de Área WICC 2019

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)
Orlando Micolini (UNC)

Computación Gráfica, Imágenes y Visualización
Martın Larrea (UNS)
Marıa J. Abasolo (UNLP-UNCPBA)
Roberto Guerrero (UNSL)

Ingeniería de Software
Pablo Fillottrani (UNSur)
Pablo Thomas (UNLP)
Fernanda Carmona (UNdeC)

Procesamiento Distribuido y Paralelo
Marcelo Naiouf (UNLP)
Fabiana Piccoli (UNSL)
Javier Balladini (UNCOMA)

Innovación en Sistemas de Software
Marcelo Estayno (UNLZ)
Guillermo Feierherd (UNTDF)
Osvaldo Spositto (UNLaM)
Gladys Dapozo (UNNE)

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 Brıa (INVAP)
Fernando Tinetti (UNLP)
Nelson Rodríguez (UNSJ)

Bases de Datos y Minería de Datos
Laura Lanzarini (UNLP)
Claudia Deco (UNR)
Norma Herrera (UNSL)

Innovación en Educación Informática
Claudia Russo (UNNOBA)
Elena Durán (UNSE)
Lucía Malbernat (UCAECE)

Seguridad Informática
Paula Venosa (UNLP)
Javier Echaiz (UNS)
Antonio Castro Lechtaller (IESE)

Jurado de Tesis de Doctorado
Marcela Printista (UNSL)
Laura de Giusti (UNLP)
Silvia Castro (UNS)
Alejandra Cechich (UNCOMA)
Horacio Kuna (UNaM)

Comité Académico

UBA – Cs. Exactas	Garbervetsky, Diego	UBA–Ingeniería	Echeverría, Adriana
UN La Plata	Pesado, Patricia	UN Sur	Rueda, Sonia
UN San Luis	Píccoli, Fabiana	UNCPBA	Aciti, Claudio
UN Comahue	Grosso, Guillermo	UN La Matanza	Spósitto, Osvaldo
UN La Pampa	Alfonso, Hugo	UN Lomas de Zamora	Estayno, Marcelo
UN Tierra del Fuego	Feierherd, Guillermo	UN Salta	Gil, Gustavo
UN Patagonia Austral	Lasso, Marta	UN San Juan	Rodríguez, Nelson
UADER	Noriega, Jorge	UN Patagonia SJB	Buckle, Carlos
UN Entre Ríos	Tugnarelli, Mónica	UN Nordeste	Dapozo, Gladys
UN Rosario	Zanarini, Dante	UN Misiones	Kuna, Horacio
UNNOBA	Russo, Claudia	UN Chilecito	Carmona, Fernanda
UN Lanús	Azcurra, Diego	UN Santiago del Estero	Duran, Elena
Esc. Sup. Ejercito	Arroyo Arzubi, Alejandro	UN Litoral	Loyarte, Horacio
UN Rio IV	Arroyo, Marcelo	UN Córdoba	Fridlender, Daniel
UN Rio Negro	Vivas, Luis	UN Hurlingham	Medrano Gustavo
UN Villa María	Prato, Laura	UN Lujan	Panessi, Walter
UN Catamarca	Poliche María Valeria	UN La Rioja	Martínez, Marcelo
UN Tres de Febrero	Oliveros, Alejandro	UN Tucumán	Luccioni, Griselda M.
UNAJ	Morales, Martın	UN Chaco Austral	Zachman Patricia
UN del Oeste	Foti, Antonio	UN de Cuyo	García Garino, Carlos
UN de Mar del Plata	Ríos, Carlos	U Morón	Padovani Hugo
UAI	De Vincenzi, Marcelo	U Belgrano	Guerci, Alberto
U Kennedy	Panizzi, Marisa	U Adventista del Plata	Bournissen Juan
UCAECE	Finocchieto, Jorge	U Palermo	Álvarez Adriana
UCA Rosario	Grieco, Sebastián	U Salvador	Zanitti, Marcelo
U Aconcagua	Giménez, Rosa	U Gastón Dachary	Ruidías, Hector Javier
UCEMA	Guglianone, Ariadna	U Austral	Cosentino, Juan Pablo
U Atlántida Argentina	Rathmann, Liliana	UCA La Plata	Bertone, Rodolfo
ITBA	Mon, Alicia	U Champagnat	Pinciroli, Fernando
UN Jujuy	Herrera Cognetta, Analía		

Comité Organizador FCEFyN UNSJ

Evangelina Sanz
Manuel Ortega
Nelson Rodríguez
María Inés Lund

Coordinadores Locales

Laura Gutiérrez	María Murazzo	Laura Aballay	Marita Masanet
Marcelo Moreno	Jorge Mercado	Silvina Migani	Maria Romagnano
Myriam Herrera	Elisa Oliva	Flavia Millan	Cristina Vera
Cintia Ferrarini	Liliana Gonzalez	Alejandra Orellana	Susana Chavez
Hector Lepez	Alejandra Malberti		

Agentes y

Sistemas Inteligentes

Aplicación de algoritmos de optimización multiobjetivo a la mezcla de distintas fuentes de minerales en el largo plazo

Oscar Daniel Chuk, Marina E. Romero, Carlos Gustavo Rodríguez Medina, Luis V. Gutiérrez, Enrique A. Núñez, Juan P. Gil

Instituto de Investigaciones Mineras. Facultad de Ingeniería.
Universidad Nacional de San Juan

Av. Libertador Gral. San Martin 1109 oeste. San Juan.
0264-4211700 (int. 389)
dchuk@unsj.edu.ar

RESUMEN

Se presentan en este trabajo estrategias basadas en optimización multiobjetivo y supervisión predictiva destinadas a lograr una mezcla óptima de las distintas fuentes de mineral que ingresan a una planta de tratamiento, para las minas Cerro Vanguardia (Santa Cruz) y Casposo (Calingasta, San Juan). El planteo tiene una concepción Multiobjetivo, debido a las diversas variables que influyen en este caso de problemas y una concepción predictiva debido a que se desea modelar el comportamiento de las variables a lo largo del tiempo.

Este algoritmo se plantea con la suficiente flexibilidad como para ser adaptados fácilmente a otras situaciones semejantes.

Palabras clave: Optimización multiobjetivo, Mezcla óptima de minerales, Programación de Mina, Supervisión predictiva.

CONTEXTO

Este trabajo tuvo su origen en el diseño por el Instituto de Investigaciones Mineras IIM de un algoritmo óptimo de mezcla de minerales para la empresa Cerro Vanguardia. El mismo tenía características estáticas, ya que resolvía el problema para una situación determinada. Este algoritmo original se ha visto mejorado con un nuevo desarrollo para la mina Casposo (Troy Resources Limited), en el que se contempla en forma óptima y dinámica no sólo la situación actual de producción sino todas las posibles etapas del proceso y vida útil de la mina hacia futuro, contemplando incluso hasta el cierre del emprendimiento.

Este desarrollo se lleva a cabo en el marco de la Convocatoria 2018-2019 de Proyectos de Investigación y Desarrollo Social PDTS de la U.N.S.J. con el título *"Programación de mina de largo plazo mediante optimización multiobjetivo predictiva"*, y cuenta con el aval del Ministerio de Minería de la Provincia de San Juan en calidad de entidad patrocinante.

1. INTRODUCCIÓN

Desde que Lerchs y Grossman (1965) publicaron su método de explotación óptima de una cantera a cielo abierto basado en programación dinámica, ha existido un creciente interés en la industria minera por los métodos que permiten maximizar o minimizar un objetivo técnico o económico. El tratamiento de este problema, conocido como "mine planning" ha ido evolucionando al ritmo del desarrollo de las técnicas computacionales de optimización en los últimos 50 años. Al respecto, si bien algunos autores como Gershon (1987) sostienen que los métodos de optimización matemática no deben dejar de lado la heurística tradicional usada por los planificadores de mina, la mayoría de las soluciones recurren a la

Programación Lineal Mixta Entera (Mixed Integer Linear Programming MILP, Askari-Nasab et al., 2010), usando habitualmente el paquete comercial de cálculo CPLEX de IBM. Otras aproximaciones han usado la técnica de "branch and cut" (Caccetta and Hill, 2003), algoritmos genéticos (Ataei and Osanloo, 2003) y soluciones derivadas del control automático tales como el control predictivo (Goodwin et al., 2006).

El presente artículo se orienta a un problema relacionado con el previamente expuesto, pero de posterior desarrollo en la academia, la mezcla óptima de diversas fuentes de mineral en la producción minera. Esto involucra la mezcla de distintas menas, como por ejemplo diferentes sectores de una mina de cielo abierto, vetas de minas subterráneas y pilas de acopio, diversas plantas de tratamiento y distintos productos. Si bien las formulaciones más simples del problema de mezcla óptima están basadas en programación lineal (Chanda and Dagdelen, 1995), el proceso de optimización involucra la toma de decisión del destino completo de lotes de mineral, lo que a su vez implica el uso de variables enteras mezcladas con otras reales, en funciones de evaluación no lineales. Se trata entonces de un Algoritmo de Programación no lineal Mixto (Mixed Integer Nonlinear Program MINLP, Blom et al., 2016).

1.1 La planificación de mezcla de fuentes de mineral como un problema multiobjetivo

El caso particular de la planificación de producción en las industrias extractivas involucra un número elevado de variables: los caudales de mineral por veta, las leyes de las especies a extraer, la vida útil de la mina, las capacidades de la/las planta/s de tratamiento, sus recuperaciones, los valores en el mercado de los minerales, los costos de extracción y tratamiento, los costos energéticos, laborales, etc.

Al mismo tiempo, existen restricciones impuestas por factores técnicos y económicos que completan el problema de optimización. Los objetivos prioritarios de optimización suelen ser la maximización del Valor Actual Neto VAN (Net Present Value NPV, Asad and Topal, 2011) o una política de ley de corte (Nesis, 2017). Pero tampoco son los únicos. Es habitual que el mismo se complemente con una consigna de producción en un determinado período y otros objetivos posibles como el agotamiento de los recursos al momento del cierre proyectado de la mina. Se tiene así un problema que contempla más de un objetivo. La resolución de un problema en el cual dos o más objetivos compiten entre sí, considerando restricciones técnico – económicas, es por definición un problema de optimización multiobjetivo.

Algunos autores como Everett (2010) resuelven esta versión multiobjetivo del problema de mezcla optimizando los objetivos iterativamente. Otros incluyen algunos objetivos como la estabilidad de los taludes en las restricciones del problema (Galić et al., 2009). Pero es difícil encontrar en la literatura abordajes estrictamente multiobjetivo del problema.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

2.1 Formulación del problema

En un planteo típico del problema, se propone la optimización de la mezcla de 10 fuentes de mineral en una operación minera metalífera de oro y plata hasta el agotamiento de las reservas, y se ha estimado que esto puede ocurrir a los N_c meses. El esquema mina-plantas a optimizar se presenta en la Figura 1.

Cada fuente de mineral tiene una previsión de reservas estimadas [tn], y está caracterizada por una ley de oro C_u [gr/tn] y de plata C_g [gr/tn], pudiendo variar ambas a lo largo de los N_c meses.

Los caudales Q_i [tn/mes] provistos por las fuentes de mineral pueden ser derivados a distintas plantas según una variable entera $d_i \in \{1, 2, 3\}$. Si $d_i = 1$ el destino es una planta de Molienda, si $d_i = 2$ se deriva a una pila de Stock de Largo Plazo SLP intermedia, y si $d_i = 3$ se destina a una planta de Lixiviación en valle HL. La pila SLP aporta a las plantas de Molienda y HL caudales Q_{SLPm} y Q_{SLPh} respectivamente. Las plantas de Molienda y HL tienen capacidades de trata-

miento Q_{scmb} y Q_{schb} [tn/mes], con recuperaciones para oro y plata R_{mu} y R_{mg}, R_{hu} y R_{hg}, respectivamente. Los datos económicos son las cotizaciones del oro *CotAu*, de la plata *CotAg*, y los costos de procesamiento de Molienda PC_m y HL PC_h.

2.2 Optimización multiobjetivo

El problema de optimización multiobjetivo se define como (Donoso y Fabregat, 2007): Encontrar el vector/es de n variables de decisión

$$\boldsymbol{x}^* = \left[x_1^* \quad x_2^* \quad \ldots \quad x_n^*\right]^T \tag{1}$$

que satisfaga las m restricciones de desigualdad

$$g_i(\boldsymbol{x}) \geq 0 \text{ para } i = 1, 2, \ldots, m \tag{2}$$

las t restricciones de igualdad

$$h_i(\boldsymbol{x}) = 0 \text{ para } i = 1, 2, \ldots, t \tag{3}$$

y optimice (maximice o minimice) el vector de q funciones de evaluación

$$\boldsymbol{f}(\boldsymbol{x}) = \left[\, f_1(\boldsymbol{x}) \;\; f_2(\boldsymbol{x}) \;\; \ldots \;\; f_q(\boldsymbol{x})\right] \tag{4}$$

Cuando hay varias funciones de evaluación $f_j(\boldsymbol{x})$ las mismas compiten entre sí, por lo cual no es posible encontrar *un* óptimo sino que el resultado es un grupo de soluciones óptimas denominado *Conjunto de Pareto* (Coello Coello, 2002).

2.3 Carácter predictivo de la optimización

El hecho de considerar el desarrollo futuro del emprendimiento en las funciones de optimización $\boldsymbol{f}(\boldsymbol{x})$, a lo largo de los N_c de vida útil, implica el diseño de las mismas con un carácter predictivo, lo cual se logra echando a andar el modelo matemático del proceso hacia adelante en el tiempo. Se toman en cuenta las predicciones de leyes minerales, costos, valores de mercado, etc., tal como se muestra en la Figura 2.

El vector $\boldsymbol{x}$ de variables de decisión es

$$\begin{aligned}\boldsymbol{x} = [&Q_1(1)\; Q_2(1) \ldots\; Q_{10}(1)\; Q_{SLPm}(1)\; Q_{SLPh}(1) \ldots \\ &Q_1(N_c)\; Q_2(N_c) \ldots\; Q_{10}(N_c)\; Q_{SLPm}(N_c)\; Q_{SLPh}(N_c) \ldots \\ &\ldots\; d_1(1)\; d_2(1) \ldots\; d_{10}(1) \ldots \\ &d_1(N_c)\; d_2(N_c) \ldots\; d_{10}(N_c)\;]\end{aligned} \tag{5}$$

El algoritmo de cálculo multiobjetivo entrega un Conjunto Óptimo de Pareto donde cada individuo del mismo es una trayectoria óptima posible de la forma (5), graficados en la parte superior de la Fig. 2 (Por simplicidad se grafican sólo los caudales Q_i, no los destinos d_i). En una segunda fase del proceso de optimización, se debe elegir alguna de estas trayectorias para aplicarla efectivamente al proceso, por un método que suele ser la menor norma euclídea al vector ideal (Chuk, 2012).

3. RESULTADOS OBTENIDOS Y ESPERADOS

El problema ha sido resuelto usando un algoritmo de optimización de colonias de hormigas (Schlueter y Munetomo, 2017).

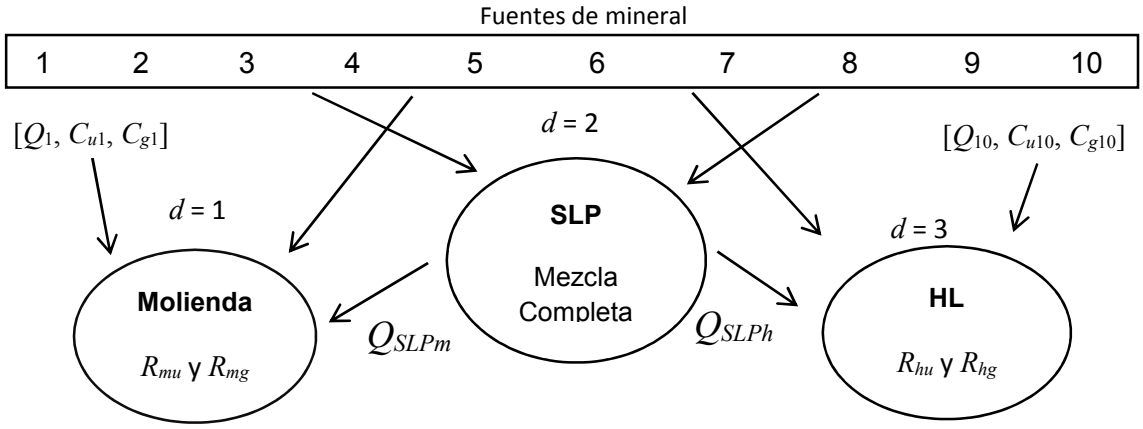


Figura 1. Esquema del conjunto mina – plantas

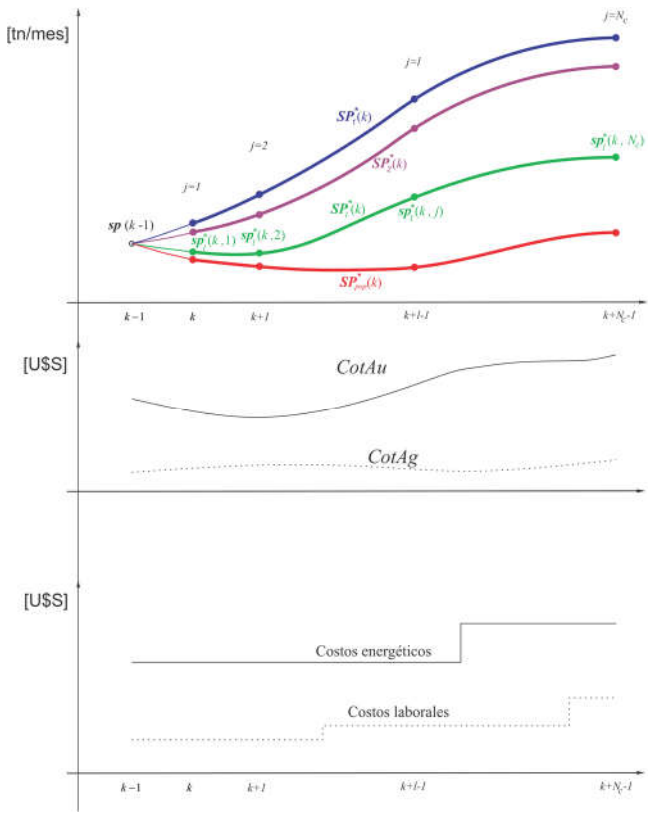


Fig. 2 Evolución futura de variables

En un ejemplo de aplicación concreto se considera un horizonte de N_c = 24 meses, leyes crecientes en orden con las fuentes de mineral, pero con una fuerte caída en las previsiones de las mismas en el mes 8 y un incremento en la cotización del oro del 13% en el mes 16. La planta de Molienda tiene una capacidad de Q_{scmb} = 50000 tn/mes y HL de Q_{scmh} =100000 tn/mes, y se le impone a la pila SLP que no descienda de 20000 tn hasta el mes 20, cuando se inicia el cierre de mina. La evolución en el tiempo se presenta en la Figura 3. El vector de 10 destinos optimizados es [3 2 2 3 3 2 1 1 1 1] para los meses 1 a 7, [3 3 2 2 2 2 3 3 2 2] para los meses 8 a 15 y [3 2 3 3 1 1 2 2 3 2] para los meses 16 a 24.
Las soluciones optimizadas obtenidas presentan un aumento del VAN de entre el 5% y 10% respecto de la solución calculada por un programador de mina entrenado. Se observa un uso máximo y constante de ambas plantas, se respeta el volumen mínimo de la pila *minVolSLP* y se asegura el agotamiento de las reservas al fin del emprendimiento.

4. FORMACIÓN DE RECURSOS HUMANOS

La Ing. Marina E. Romero lleva adelante sus tesis de doctorado en Ingeniería de Procesamiento de Minerales en la temática, con el título "Planificación de mina de largo plazo basada en optimización multiobjetivo y supervisión predictiva".

5. BIBLIOGRAFIA

Asad, M.W.A. and Topal, E. (2011) 'Net present value maximization model for optimum cut-off grade policy of open pit mining operations'*„J. S. Afr. Inst. Min. Metall*, 111 (11), 741-750.

Askari-Nasab, H., Awuah-Offei, K. and H. Eivazy (2010) 'Large-scale open pit production scheduling using Mixed Integer Linear Programming', International Journal of Mining and Mineral Engineering, 2 (3), 185-214.

Ataei, M., and Osanloo, M. (2003) 'Using a combination of genetic algorithm and the grid search method to determine optimum cutoff grades of multiple metal deposits', International Journal of Surface Mining, Reclamation and Environment, 18 (1), 60–78.

Blom, M.L., Pearce, A.R. and Stuckey, P.J. (2016) 'A Decomposition-Based Algorithm for the Scheduling of Open-Pit Networks Over Multiple Time Periods'. *Management Science*. http://dx.doi.org/10.1287/mnsc.2015.2284

Caccetta, L., and Hill, S. P. (2003) 'An application of branch and cut to open pit mine scheduling'. Journal of Global Optimization, 27 (2–3), 349–365.

Chanda, E.K.C. and K. Dagdelen (1995) 'Optimal blending of mine production using goal programming and interactive graphics systems'. International Journal of Surface Mining, Reclamation and Environment, 9 (4), 203-208.

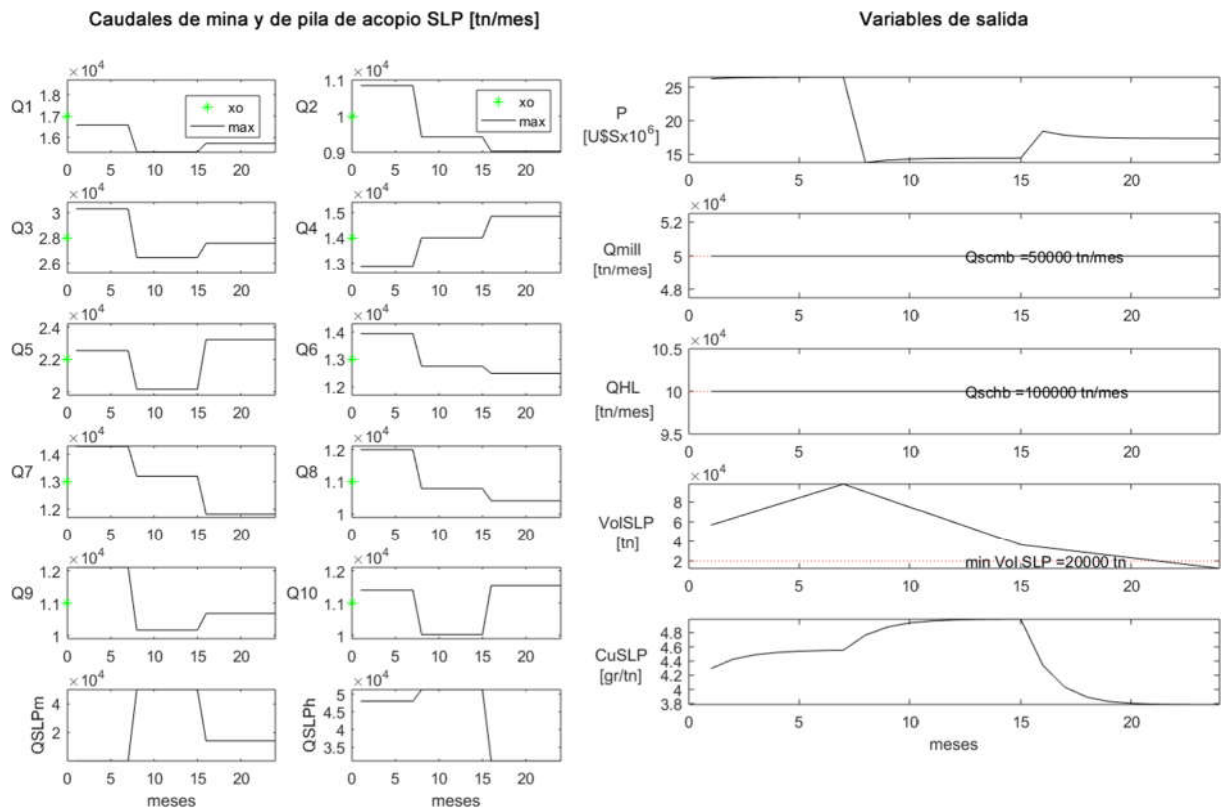


Figura 3: Variables de decisión y salidas optimizadas

Chuk, O.D. (2012) Supervisión predictiva con optimización multiobjetivo. Editorial EFU.

Coello Coello, C.A., Van Veldhuizen, D.A. and Lamont, G.B. (2002) Evolutionary Algorithms for Solving Multi-Objective Problems. Kluwer Academic Publishers. New York.

Donoso, Y. and Fabregat, R. (2007) Multi-Objective Optimization in Computer Networks Using Metaheuristics. Auerbach Publications, Boca Ratón.

Everett, J.E. (2010) 'Simulation Modeling of an Iron Ore Operation to Enable Informed Planning'. Interdisciplinary Journal of Information, Knowledge, and Management, 5, 101-114.

Galić, I., Jakonvić, B., Mrakovčić, I., (2009)'An another way for open pit mine design optimization – Floating slopes method', Rudarsko-geološko-naftni zbornik, 21, 103-111.

Gershon, M. (1987) 'Heuristic approaches for mine planning and production scheduling', International Journal of Mining and Geological Engineering, 5 (1), 1-13.

Goodwin, G.C., Serona, M.M., and Menabdeb, M. (2006) 'Receding horizon control applied to optimal mine planning'. Automatica, 42 (8), 1337 – 1342.

Lerchs, H. and Grossman, I. F. (1965) 'Optimum design of open-pit mines', The Canadian mining and metallurgical bulletin, 58 (633), pp. 47-54.

Nesis, V.N. (2017) 'Cutoff evaluation practice in short-term gold mine planning'. Gornyi Zhurnal, 4, 43-48.

Schlueter, M. and M. Munetomo, (2017) 'MIDACO parallelization scalability on 200 MINLP benchmarks'. Journal of Artificial Intelligence and Soft Computing, 7 (3), 171-181.

Aplicación de Metaheurísticas para la Resolución de Problemas de Optimización Dinámica

Juan José Barbero, Martín Tamagusku,
Natalia Stark, Hugo Alfonso[1],
Carlos Bermudez, Gabriela Minetti[1], Carolina Salto[1]
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. Actualmente estamos abocados al estudio del problema de diseño de redes de distribución de agua, problema al que se enfrentan las ciudades modernas con restricciones cambiantes. Este problema lo abordamos mediante el uso de metaheurísticas como *Simulated Annealing* y *Cuckoo Search* con resultados muy prometedores.

Además, hemos empezado a investigar algoritmos eficientes que puedan dar respuesta en tiempo real a diversos tipos de problemas que a su vez van cambiando dinámicamente sus restricciones. Concretamente estamos introduciendonos en el uso de los algoritmos conocidos con el nombre de "Evolución Diferencial", los cuales son flexibles para adaptar el proceso de búsqueda frente a cambios de restricciones en la función objetivo conforme avance el tiempo. Estos cambios, muchas veces provocan que soluciones factibles dejen de serlo. Una alternativa frecuentemente usada, frente a estas situaciones, es la de aplicar un mecanismo de reparación a tales soluciones del problema.

Palabras claves: Metaheurísticas, Optimización, Diseño de Red de Distribución de Agua, Simulated Annealing, Evolución Diferencial.

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.

Uno de los principales frentes de trabajo en el ámbito de las Ciencias de la Computación ha sido tradicionalmente el diseño de algoritmos cada vez más eficientes para la solución de problemas, tanto de optimización como de búsqueda. En este dominio, el objetivo consiste en obtener algoritmos nuevos que den solución al problema y que necesiten un esfuerzo computacional más pequeño que los algoritmos existentes, así como caracterizar su comportamiento para las clases de problemas que demanda la comunidad científica e industrial en general.

La investigación de algoritmos tanto exactos como heurísticos para resolver problemas de optimización tiene una vigencia inusualmente importante en estos días, ya que nos enfrentamos a nuevos problemas de ingeniería al mismo tiempo que contamos con nuevos recursos computaciones tales como nuevos tipos de máquinas, redes y entornos como Internet.

Actualmente nos enfrentamos a problemas de alta complejidad en el que intervienen varias variables con un conjunto de restricciones definidas sobre ellas, muchas veces contrapuestas, que deben ser consideradas para evaluar la factibilidad de la solución aportada, las cuales en muchos casos cambian dinámicamente [1]. 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.

En una de las líneas 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.

Otra línea de investigación se relaciona al uso de la Optimización Evolutiva en Espacios Dinámicos [2], para lo cual se utiliza un tipo de algoritmo heurístico que se conoce con el nombre de "Evolución Diferencial" e inicialmente se experimenta con un problema que cambia sus restricciones en forma dinámica [3] y es complejo de resolver por el landscape que presenta, debiendo sortear zonas de soluciones no factibles. Particularmente se fijó como hipótesis de trabajo diseñar un mecanismo de reparación de soluciones no factibles para reemplazarla por otras [4] localizadas en un contexto de vecindario delimitado por un distancia mínima.

Desarrollo

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

Optimización del diseño de redes de distribución de agua (Water Distribution Network Design Optimization WDND)

Se requiere un método de solución eficaz que sea confiable y fácil de usar para la optimización de las WDND, 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 diseña dos variantes metaheurísticas que resuelven este problema, SA y CS.

Dada las características propias de este tipo de redes, una vez que SA o CS arman una determinada solución, la prueban y evalúan por medio del simulador EPANET 2.0 [5]. Este simulador, además, resuelve todas las ecuaciones hidráulicas de forma externa. Los problemas que se utilizan para efectuar los experimentos son de período simple y multi período, en donde el patrón de demanda varía con el tiempo.

Optimización de Problemas con Restricciones Dinámicas

Un Problema de Optimización con Restricciones Dinámicas (DCOP - Dynamic Constrained Optimization Problem) puede ser visto como un problema donde el espacio de búsqueda y la región de soluciones factibles cambia a lo largo del tiempo [6] [7]. Se han definido 4 variantes de DCOPs en función de si la función objetivo y las restricciones se mantienen estáticas o dinámicas a lo largo del tiempo [2]. Este problema lo hemos abordado con Algoritmos Evolutivos Diferenciales, ellos son identificados con la sigla DE, proveniente de su denominación anglosajona ***D***ifferential ***E***volution (DE) [8]. Se trata de un algoritmo de búsqueda estocástica que opera con una población de soluciones denominadas vectores. La población está representada como: $\boldsymbol{x}_{i,G}, i = 1 \ldots NP$, donde $\boldsymbol{x}_{i,G}$ representa el vector i en la generación G, y NP es el tamaño de la población de soluciones con las que realiza la evolución usando operadores específicos de mutación y crossover.

Una variante propuesta por Mezura Montes et al [9] para mejorar el algoritmo original introduce un mecanismo de reparación de soluciones no factibles usando un método de reparación basado sobre el operador de mutación diferencial que chequea su factibilidad y en caso de no serlo, se vuelve a aplicar el mencionado mecanismo mientras no supere una cantidad $Limite_Reparacion$ o bien sea factible. El vector obtenido luego de ese proceso de reparación, sea o no factible, pasará a ser el vector obtenido para la próxima generación. El método de reparación sólo evalúa la factibilidad del vector en función de si cumple las restricciones, no por su calidad con lo cual se mantiene la diversidad genética.

A partir de ello estamos trabajando en la puesta a punto de una variante del método de reparación que consiste en buscar en el vecindario, circundante a la solución no factible, una solución factible para reemplazarla. La idea central del método es hacer prevalecer la localización del vector mutante generado y centrar la búsqueda de una solución vecina próxima a ella. La proximidad se fija con un parámetro δ que indica el desplazamiento en cada uno de las dimensiones, estimando sea el método de reparación menos disruptivo en el proceso de búsqueda.

Resultados Obtenidos

A continuación se detallan los resultados obtenidos para la primera línea de investigación mencionada, ya que el desarrollo de las variantes del mecanismo de reparación para el Problema de Optimización con Restricciones Dinámicas se encuentra en la fase de recolección de resultados.

Para resolver el problema WDND experimentamos con dos metaheurísticas, SA y CS, con respecto al CS fue desestimado, luego de efectuar diversos experimentos, ya que obtuvimos resultados de pobre calidad, sobre todo con el tiempo de resolución del algoritmo, este motivo está justificado por ser poblacional y la cantidad de evaluaciones contra los restantes algoritmos comparados es muy superior. Por otra parte adaptamos un algoritmo SA al cual le incorporamos diversos mecanismos para intensificar la búsqueda local por cada paso. Denominamos nuestra propuesta Hybrid Simulated Annealing(HSA) y consiste en tomar en cada iteración, una solución factible y aplicarle el método de búsqueda local MP-GRASP [10] que reemplazará a la solución inicial en función de un método de selección voraz. Luego se genera otra solución a partir de la inicial aplicando un operador de mutación para obtener una solución vecina y poder explorar otras áreas del espacio de búsqueda, esta solución puede ser aceptada en función de una probabilidad de Boltzmann, que depende de la temperatura actual. De esta forma con temperaturas altas se favorece la exploración, en contraste con una temperatura baja en donde se hace una búsqueda intensiva en alguna región prometedora. La temperatura se actualiza con el método de enfriado proporcional [11], aplicada luego de un número de iteraciones determinadas por la longitud de la cadena de Markov. Nuestros resultados fueron comparados con los resultados obtenidos en [10] y pudimos mejorar los resultados en un 60 % de las instancias, además de obtener una convergencia mucho más rápida hacia las mejores soluciones.

En la actualidad se están probando diferentes variantes para mejorar los resultados de este problema, sobre todo centrándonos en el aspecto dinámico, ya que es más representativo de las necesidades reales del campo de aplicación de estas redes. Una metaheurística como la evolución diferencial [12] reúne estas características y nos encontramos en la actualidad recabando información para determinar la factibilidad de aplicar esta técnica.

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 tesinas 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] "Immune generalized differential evolution for dynamic multi-objective environments: An empirical study," *Knowledge-Based Systems*, vol. 142, pp. 192 – 219, 2018.

[2] M.-Y. Ameca-Alducin, E. Mezura-Montes, and N. Cruz-Ramírez, "Dynamic differential evolution with combined variants and a repair method to solve dynamic constrained optimization problems: an empirical study," *Soft Computing*, vol. 22, no. 2, pp. 541–570, Jan 2018. [Online]. Available: https://doi.org/10.1007/s00500-016-2353-1

[3] K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms*, ser. Wiley Interscience Series in Systems and Optimization. Wiley, 2001. [Online]. Available: https://books.google.com.ar/books?id=OSTn4GSy2uQC

[4] M. Gen and R. Cheng, *Genetic Algorithms and Engineering Design*, ser. A Wiley Interscience publication. Wiley, 1997. [Online]. Available: https://books.google.com.ar/books?id=MCHCaJAHFJAC

[5] *Users Manual*, 2000.

[6] T. T. Nguyen and X. Yao, "Continuous dynamic constrained optimization—the challenges," *IEEE Transactions on Evolutionary Computation*, vol. 16, no. 6, pp. 769–786, Dec 2012.

[7] E. Juárez-Castillo, N. Pérez-Castro, and E. Mezura-Montes, "An improved centroid-based boundary constraint-handling method in differential evolution for constrained optimization," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 31, no. 11, p. 1759023, 2017. [Online]. Available: https://doi.org/10.1142/S0218001417590236

[8] K. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*.

[9] E. Mezura-Montes, M. E. Miranda-Varela, and R. del Carmen Gómez-Ramón, "Differential evolution in constrained numerical optimization: An empirical study," *Inf. Sci.*, vol. 180, pp. 4223–4262, 2010.

[10] A. De Corte and K. Sörensen, "An iterated local search algorithm for water distribution network design optimization," *Network*, vol. 67, no. 3, pp. 187–198, May 2016.

[11] S. Kirkpatrick, C. G. Jr, and M. Vecchi, "Optimization by simulated annealing," *Science*, no. 220, pp. 671–680, 1983.

[12] P. K. Storn, R., "Differential evolution—a simple and efficient adaptive scheme for global optimization over continuous space," *Technical Report TR-95-012*, 1995.

Aplicación de minería de datos en la extracción de perfiles de deserción estudiantil sobre 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

El presente proyecto plantea como objetivo la identificación de los diferentes factores que influyen en la deserción estudiantil, con el fin de caracterizar y extraer los perfiles de los estudiantes desertores. Se espera así, generar un conjunto de indicadores de detección temprana de problemáticas del aprendizaje y dificultades en el trayecto inicial de formación académica.

Para tal fin, 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, tomando como caso de estudio a los ingresantes y aspirantes a las carreras brindadas por la Universidad Tecnológica Nacional Facultad Regional La Plata.

Palabras clave: Deserción Estudiantil, Deserción Universitaria, Minería de Datos Educacional, Indicadores de Deserción, Indicadores de Riesgo de Deserción.

CONTEXTO

En Argentina, el fenómeno de la deserción estudiantil universitaria cobra especial interés en las carreras de Ingeniería. 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 (PEFI, 2012). Con el fin de aumentar los egresados en estos campos, las carreras de Ingeniería se encuentran enmarcadas dentro del Plan Estratégico de Formación de Ingenieros 2012-2016.

El plan tiene tres ejes principales, el primero de ellos se centra en los Proyectos de Mejoramiento de Indicadores Académicos, y reúne distintas estrategias con las que intenta incrementar la cantidad de graduados en ingeniería en un 50% en 2016 y en un 100% en 2021, con relación al año 2009.

En concordancia con este plan y para atender a las demandas sociales y propias de la institución, la UTN-FRLP se encuentra actualmente trabajando en el Proyecto de Investigación y Desarrollo (PID): "Estudio Sistemático de Deserción Estudiantil Universitaria (ESDEU)", el cual tiene como objetivo final el desarrollo e implementación un sistema de gestión tutorial, con el fin de elevar la tasa de graduación en cada una de las especialidades que brinda (Del Giorgio & Istvan, 2016); dentro del cual nace la línea de investigación presentada en este trabajo.

El proyecto ESDEU se encuentra homologado y financiado por la Universidad Tecnológica Nacional, UTN con el código TEINNLP0003786.

1. INTRODUCCIÓN

Con el fin de poder definir los perfiles de los alumnos desertantes, la UTN FRLP utiliza los datos registrados en la base de datos académica,

perteneciente al sistema SysAcad provisto por Rectorado UTN, la cual dispone de un número importante de atributos personales y académicos; que actualmente se complementan principalmente con los datos almacenados en una fuente de datos externa, que contiene los registros de las Encuestas PEFI llevadas a cabo en el período 2013/2017. Estas encuestas, respondían al pedido de la Secretaría de Políticas Universitarias, enmarcadas dentro del Plan Estratégico de Formación de Ingenieros (PEFI, 2012) y se completaban a través de una plataforma web, de forma obligatoria por los mismos estudiantes. Recogían información complementaria, registrándose datos personales, laborales, familiares y económicos, como ser: Horas de Trabajo, Relación del Trabajo con la Carrera, Situación Económica de los Padres, Nivel de Instrucción de los Padres y Becas, entre otros.

De forma conjunta los datos académicos propios de la regional y las encuestas PEFI, conforman una base de atributos relevantes cuya disponibilidad constituye un importante activo en la institución, en la medida en que son transformados en información de interés, utilizando técnicas y métodos para definir el perfil de los estudiantes.

Para enriquecer la investigación se utilizan como métodos e instrumentos de apoyo las encuestas antes mencionadas y la técnica de Minería de Datos, cuyo fin es trabajar en la detección de relaciones entre variables de manera automática y patrones de comportamiento que indican las posibles causas de deserción estudiantil.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La Minería de Datos formalmente reúne un conjunto de técnicas capaces de modelizar y resumir la información, facilitando su comprensión y ayudando a la toma de decisiones. Surge como un campo de la Estadística y las Ciencias de la Computación dando respuesta al intento de descubrir patrones en conjuntos inmanejables de datos.

Esta técnica, basada en la extracción de conocimiento, está principalmente relacionada con el proceso de descubrimiento conocido como “Descubrimiento de Conocimiento en Bases de Datos” (KDD, por sus siglas en inglés), el cual se refiere al proceso no-trivial de descubrir conocimiento e información potencialmente útil dentro de los datos contenidos en repositorios de información (Baldino & Lanzarini, 2016).

Las etapas correspondientes al proceso KDD que sigue son: selección de datos, preparación, limpieza, integración y transformación de datos, minería de datos (obtención de reglas de clasificación), interpretación y evaluación (confrontación y validación de modelos) (Istvan & Chong Arias, 2016).

En el caso puntual del fenómeno de deserción estudiantil, caracterizar a los estudiantes de una institución académica aporta información no trivial y de utilidad para la gestión, posibilitando el establecimiento de políticas tendientes a mejorar el desempeño académico de los alumnos; lo cual, redunda directamente en la reducción de la deserción universitaria (Lanzarini, 2015) promoviendo asimismo aportes sustanciales en el área de Tecnología Aplicada a la Educación.

3. RESULTADOS OBTENIDOS / ESPERADOS

Hasta el momento se han obtenidos perfiles para el grupo de estudiantes ingresantes a la institución en el período 2013/2015 (Istvan, 2016; Falco, 2017) y confrontados los mismos con el modelo de indicadores resultante para los alumnos en el período de ingreso 2016/2017 (Istvan & Lasagna, 2018).

Se ha logrado elaborar un estudio de las variables socioeconómicas y académicas que caracterizan a los ingresantes de la UTN Facultad Regional La Plata. Se puede afirmar

como conclusión, que los modelos convergen hacia los siguientes niveles de incidencia de los indicadores:
Entre los factores de incidencia alta se destaca: el Nivel de Instrucción de la Madre, la Situación Económica Familiar, el Aspecto Laboral y la Relación del Trabajo con la Especialidad.
Como factores de incidencia media se observa el Nivel de Instrucción del Padre y la Cantidad de Familiares a Cargo.
En menor medida, como factor de incidencia baja se presenta la Distancia a la Institución.
Se plantea como trabajo futuro, ampliar los casos de estudio incorporando a los alumnos ingresantes en el período 2018/2019.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de desarrollo del proyecto trabaja desde el año 2014 en diversas áreas relacionadas con la deserción y caracterización de los estudiantes de la institución.
Actualmente forman parte del equipo un Director, un Coordinador, un Docente Investigador del Laboratorio de Ingeniería en Sistemas de Información (LINES), un Docente 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, un graduado y dos alumnos becarios de investigación.
Cuenta con un plan de tesis de Maestría en Tecnología Informática Aplicada en Educación en curso y dos desarrollos de Prácticas Supervisadas (PS) de los estudiantes, necesarias para la obtención del título de grado de la carrera de Ingeniería en Sistemas de Información.

5. BIBLIOGRAFÍA

[1] Baldino, Guillermo; Lanzarini, Laura; Charnelli, María Emilia (2016); “Análisis del avance académico de alumnos universitarios. Un estudio comparativo entre la UTN-FRLP y la UNLP”; XI Congreso de Tecnología en Educación y Educación en Tecnología; ISBN: 978-987-3977-30-5; p. 589-596.

[2] Del Giorgio Paula; Ressel Ismain; Tutores: Istvan Romina, Falco Mariana, Chong Arias Carlos (2016); “Herramienta para la detección temprana de deserción estudiantil para la UTN FRLP - Diseño preliminar”, CONAIISI 4to Congreso Nacional de Ingeniería Informática, 17 y 18 de noviembre de 2016, UCASAL Universidad Católica de Salta.

[3] Falco Mariana, Istvan Romina, Antonini Sergio (2016); “Deserción Universitaria: Validación de Patrones Relacionales en los Alumnos de Primer Año”, CONAIISI 4to Congreso Nacional de Ingeniería Informática, 17 y 18 de noviembre de 2016, UCASAL Universidad Católica de Salta.

[4] Himmel Erika (2002); “Modelos de Análisis de la Deserción Estudiantil en la Educación Superior”; Revista Calidad de la Educación, Chile, 2002.

[5] Istvan Romina; Chong Arias Carlos; Antonini Sergio (2016); “Sistemas de indicadores de riesgo de deserción para la UTN-FRLP. Análisis sobre alumnos ingresantes en los años 2013/2015”, CYTAL 2016 – 7° Jornadas de Ciencia y Tecnología, 12 al 14 de octubre de 2016, UTN FRVM.

[6] Istvan Romina; Lasagna Valeria (2018); “Detección Temprana de Deserción Estudiantil para Aspirantes e Ingresantes de la UTN FRLP”, CYTAL 2018 – 8° Jornadas de Ciencia y Tecnología, 12 al 14 de septiembre de 2018, UTN FRVM.

[7] Lanzarini L, Charnelli E, Baldino G., Díaz J. (2015); “Selección de atributos representativos del avance académico de los alumnos universitarios usando técnicas de visualización. Un caso de estudio”; Revista TE&ET; no. 15. ISSN: 1850-9959. p. 42-50. 2015.

[8] PEFI Plan Estratégico de Formación de Ingenieros 2012/2016; Secretaría de Políticas Universitarias, Ministerio de Educación, Presidencia de la Nación Argentina; http://pefi.siu.edu.ar/, 2012. Accedido en abril de 2018.

[9] Tinto V. (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.

Aplicación de una arquitectura de red neuronal para el monitoreo de carga por métodos no invasivos (NILM) utilizando ciclos de activación de artefactos eléctricos en el entrenamiento

Diego Cocconi*, Rebeca Yuan*, Micaela Mulassano*, Diego Ferreyra*

Departamento de Ingeniería en Sistemas de Información / Electromecánica
Universidad Tecnológica Nacional (UTN), Facultad Regional San Francisco
San Francisco (2400), Córdoba, Argentina
*{dcocconi, ryuan, mmulassano, dferreyra}@sanfrancisco.utn.edu.ar

Resumen

Con el objetivo de lograr identificar artefactos eléctricos utilizando redes neuronales a partir de una medida total de consumo de energía (técnica conocida como NILM, del inglés *Non-Intrusive Load Monitoring*), en el presente trabajo se plantea la evaluación de dos tipos de redes neuronales capaces de realizar tal tarea, contando como ejemplos de entrenamiento válidos para el aprendizaje con ciclos de activación de diferentes artefactos que ya fueron identificados por un algoritmo de detención desarrollado en trabajos anteriores.

Palabras Claves: NILM – consumo de energía – ciclos de activación – aprendizaje automático – redes neuronales

CONTEXTO

El presente trabajo forma parte del proyecto de investigación I+D UTN CCUTNSF0004881 "*Monitoreo domiciliario utilizando redes neuronales a partir de una medición de energía totalizada (NILM)*". El mismo está homologado como proyecto de investigación y desarrollo por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional (UTN). En el marco de dicho proyecto se propone implementar NILM mediante *aprendizaje automático* (del inglés *machine learning*) para el análisis del consumo de energía en los hogares argentinos, a través de la utilización de redes neuronales artificiales.

1. INTRODUCCIÓN

Para identificar el consumo eléctrico de los diferentes artefactos en una vivienda existen dos formas de realizar la medición: [1-2]: (1) mediante la utilización de medidores independientes para cada artefacto y/o algún equipo central que recolecte la información; (2) la aplicación de *Non-Intrusive Load Monitoring* (NILM) [3] o *Non-Intrusive (Appliance) Load Monitoring* (NIALM o NALM) [3-4], que podría traducirse como "*monitoreo de carga no invasivo*" o "*monitoreo de carga por métodos no invasivos*". NILM es una técnica computacional que, a partir de una medición total de consumo de energía, logra identificar los artefactos eléctricos individuales que se encuentran consumiendo la misma [5]; medir cada artefacto conlleva a un método más exacto que NILM en términos de resultados obtenidos, pero las desventajas como costos elevados, múltiples configuraciones de sensores y complejidad en la instalación favorecen el uso del NILM [6]. Un enfoque comúnmente utilizado para implementar NILM involucra las siguientes etapas: (1) *adquisición de datos*; (2) *extracción de*

patrones; e (3) *inferencia o aprendizaje automático* [6].

El proyecto que se describe en este trabajo pretende abarcar las tres etapas. Para la primera parte, se está trabajando en una fase de prueba para la adquisición de datos y poder obtener muestras de diversos consumos en nuestra Facultad. La segunda etapa, que consiste en la extracción de patrones, ya ha sido implementada y ha producido los resultados esperados, descritos en un trabajo previo: [7]. La misma ha implicado identificar los datos de muestras de diversos artefactos individuales (que podrían ser ondas de voltaje y corriente, o potencia, como se observa en la Figura 1, la cual muestra las ondas de potencia instantánea de dos artefactos comunes: un lavarropas en la parte superior y una pava eléctrica en la parte inferior), con el objetivo de extraer métricas de potencia (potencia reactiva, potencia activa, potencia media, etc.), así como eventos de transición de estados (estados "encendido/apagado" u "ON/OFF"). Esta etapa tiene un impacto directo en siguiente fase de aprendizaje, pues permite obtener ejemplos de calidad para el entrenamiento de la técnica de aprendizaje automático. Si el conjunto de ejemplos no contiene buenos datos o datos de calidad, el aprendizaje posterior se verá afectado. Ante la falta de datos de muestreo locales, se ha decidido utilizar como punto de partida los datos proporcionados por el repositorio UK-DALE (su parte de baja frecuencia, con datos muestreados cada 6 segundos). El objetivo fue encontrar un algoritmo capaz de identificar los *ciclos de activación* (períodos durante los cuales los artefactos se encuentran consumiendo energía) de los distintos artefactos eléctricos medidos en forma individual, sin considerar ruidos, perturbaciones y señales de anomalía. Por ejemplo, en la Figura 2 se puede apreciar un ciclo de activación de un lavarropas, en medio de dos grupos de interferencias, que deben ignorarse (en t_1, entre t_2 y t_3, y en t_8 y t_9); los retornos a cero también constituyen situaciones que deben ignorarse, pues pueden confundirse con el final de un ciclo de activación y el comienzo de otro nuevo (t_5 y t_6). La Figura 3 muestra señales reales de este tipo de anomalías.

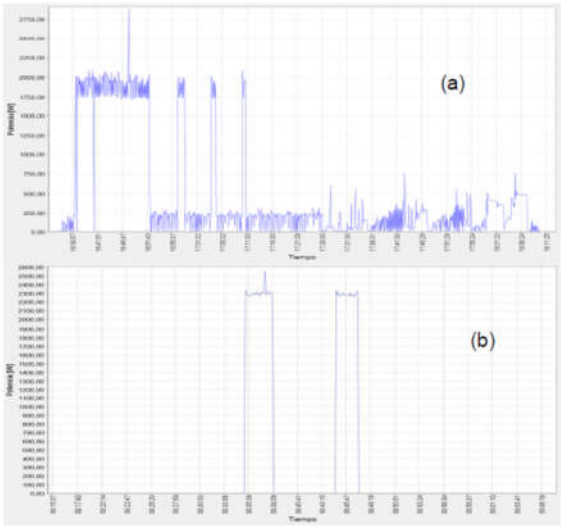


Figura 1: Consumo de dos artefactos eléctricos a través del tiempo, en baja frecuencia (6 segundos), usando datos de origen del repositorio UK-DALE[1]. (a) Lavarropas. (b) Pava eléctrica.

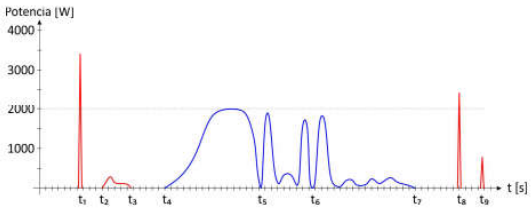


Figura 2: Onda de consumo de ejemplo, mostrando posibles perturbaciones [7].

Para la correcta extracción de los ciclos de activación de los artefactos, el grupo de investigación desarrolló un algoritmo de detección, el cual fue implementado y validado con los datos del repositorio

[1] http://jack-kelly.com/data/

seleccionado, donde se parte de lecturas de potencia y tiempo de los artefactos (considerando baja frecuencia), y determina en tres pasos si una medición dada contiene (1) el inicio de un ciclo, (2) el final del mismo y (3) valores acordes entre inicio y fin. Es mediante la interacción de estas tres acciones que el algoritmo permite extraer el ciclo de activación completo de un artefacto.

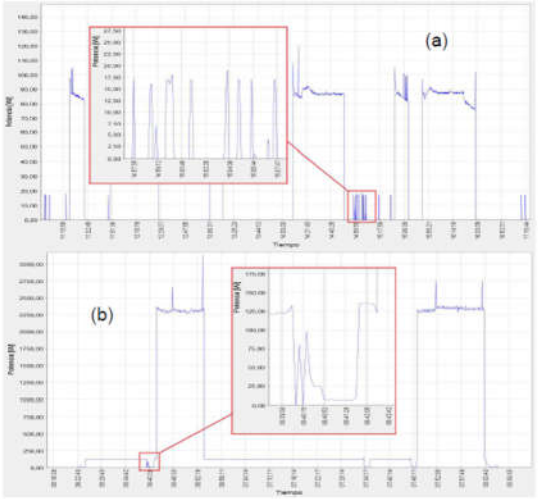


Figura 3: Aspectos a tener en cuenta durante la detección de ciclos de activación (datos del repositorio UK-DALE). (a) Perturbaciones (heladera). (b) Retornos a cero (lavavajillas).

Diseñado el algoritmo, se procedió a la implementación y prueba del mismo. Para ello, primeramente, se cargaron los valores del repositorio UK-DALE en una base de datos MySQL, debido a que los datos originales se encontraban en archivos de texto plano. Una vez transferidos a la base de datos, los valores pueden buscarse con mayor facilidad, para estudiar las mediciones de un día determinado, por ejemplo. Mediante una aplicación Java, la cual implementa el algoritmo, es posible: (1) seleccionar un período para obtener la medición (por ahora, es posible seleccionar períodos equivalentes a un día completo, por lo que la granularidad es de un día); (2) especificar el artefacto a evaluar para encontrar los ciclos de activación; y (3) configurar los parámetros del algoritmo. Una vez especificada esta información, es posible graficar la medición y los ciclos de activación detectados, para determinar visualmente si dichos ciclos de activación son correctos. Los ciclos identificados brindan información acerca de los mismos: (1) su potencia media; (2) su potencia máxima; y (3) su duración [7]. La Figura 4 muestra dicho gráfico, con tres ciclos detectados para el caso de un lavarropas.

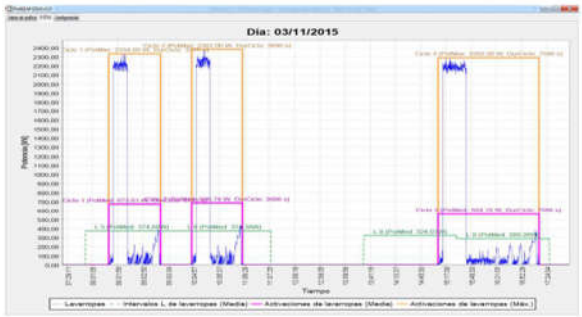


Figura 4: Vista del gráfico que muestra los resultados del algoritmo.

Mediante esta introducción se ha descripto el trabajo realizado en el transcurso del primer año del proyecto de investigación, se pretende ahora comenzar la última etapa del proyecto, que consiste en desarrollar la fase de aprendizaje automático mediante la evaluación y selección de una arquitectura apropiada de red neuronal y la integración de lo implementado en las tres fases.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

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 de energía eléctrica.

- Agentes y sistemas inteligentes.

3. RESULTADOS ESPERADOS

Tal como quedó expuesto, se lograron identificar los ciclos de activación de varios artefactos eléctricos y depurar la base de datos de muestras con el objetivo de contar con ejemplos consistentes para el entrenamiento de las herramientas de aprendizaje automático. La aplicación de aprendizaje automático busca el detectar el funcionamiento de cada artefacto eléctrico, con el objetivo a futuro de detección de distintas anomalías, como excesos injustificados de consumo o mal uso del artefacto.

Al contar con datos depurados, el tipo de aprendizaje que se pretende realizar será del tipo supervisado; esto significa que los datos deben encontrarse etiquetados. Se analizarán distintas estructuras de redes neuronales para el aprendizaje, pero se hará foco en: (1) *redes recurrentes* y (2) *redes autoencoder*. Las redes recurrentes trabajan con datos secuenciales (fundamental para los tipos de datos que se manejan) [8] y constituyen redes de *retro-propagación* (o del inglés, *backpropagation*). Esta retro-propagación le permite a la red generar y detectar patrones variantes en el tiempo. Dentro de las redes recurrentes, se analizará la *red Elman*, que trabaja con dos funciones de activación (*tansigmosoidal* en sus capas ocultas y una función lineal en su salida), lo que permite aproximar cualquier función (considerando el comportamiento discontinuo del funcionamiento de los artefactos) con cierta exactitud en base a la cantidad de neuronas en su capa oculta. Por otro lado, y buscando un tipo de aprendizaje relacionado con imágenes, se analizarán redes autoencoder, que pertenecen al conjunto de *redes convolucionales*, diseñadas para descubrir características de imágenes [9], por medio de operaciones no lineales. En este tipo de estructura la comunicación es hacia delante, pero no todas las neuronas de una capa oculta reciben la información de todas las neuronas de la capa anterior, lo que conduce a reducir el número de capas y de operaciones entre neuronas.

Con el estudio, implementación y comparación de estas dos estructuras, se buscará determinar cuál es la óptima para dar solución al problema.

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 realizando su carrera de Doctorado en Ingeniería, con mención en Sistemas de Información (en la UTN Facultad Regional Santa Fe), otro ha finalizado recientemente su Doctorado en Ciencias de la Ingeniería, con mención Ingeniería Eléctrica (en la Universidad Nacional de Río Cuarto) y dos se encuentran desarrollando sus tesis para las Maestrías en Ingeniería de Software y Calidad de Software (en la UTN Facultad Regional San Francisco), quienes a su vez dictan la cátedra Inteligencia Artificial (Ingeniería en Sistemas de Información) en la UTN Facultad Regional San Francisco.

Se prevé la capacitación y formación de recursos humanos, a través de cursos de actualización y posgrado en el área de estudio; la transferencia de conocimiento y resultados; y la posibilidad de brindar charlas informativas del desarrollo e implementación del proyecto a distintos sectores de la industria interesados y cátedras afines a la investigación.

5.BIBLIOGRAFÍA

[1] Dong, M., Meira, P. C., Xu, W. y Freitas, W. (2012). “An event window based load monitoring technique for smart meters”. *IEEE transactions on smart grid*, 3(2), 787-796.

[2] Ridi, A., Gisler, C. y Hennebert, J. (2014, Agosto). “A survey on intrusive load monitoring for appliance recognition”. *Pattern Recognition (ICPR), 2014 22nd International Conference on* (pp. 3702-3707). IEEE.

[3] Bernard, T. y Marx, M. (2016, Mayo). “Unsupervised learning algorithm using multiple electrical low and high frequency features for the task of load disaggregation”. *Proceedings of the 3rd International Workshop on NILM*, Vancouver, BC, Canada (pp. 14-15).

[4] Hart, G. W. (1992). “Nonintrusive appliance load monitoring”. *Proceedings of the IEEE*, 80(12), 1870-1891.

[5] Kelly, J. y Knottenbelt, W. (2015, Noviembre). “Neural nilm: Deep neural networks applied to energy disaggregation”. *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments* (pp. 55-64). ACM.

[6] 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.

[7] Cocconi, D., Yuan, R., Mulassano, M., Ferrero, N., Beltramone, M. y Biasco A. (2018). “Diseño de un algoritmo para la detección eficaz de ciclos de activación en la aplicación de monitoreo de carga por métodos no invasivos (NILM)” *6to Congreso Nacional de Ingeniería en Informática / Sistemas de Información, Universidad CAECE*. Mar del Plata, Buenos Aires, Argentina.

[8] Contreras, W., Arichávala, M., y Jeréz, C. (2018). "Determinación de la presión máxima de compresión de un motor de encendido provocado basado en una red neuronal artificial recurrente". *Ingenius*.

[9] Pusiol, P. D. (2014). "Redes Convolucionales en Compresión de Escenas". FAMAF- UNC, Córdoba.

Aprendizaje automático y análisis de autoría. Investigaciones con textos periodísticos argentinos

Mercado V[1]., Villagra A.[1], Errecalde M.[1,2]

[1]Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada (ITA) - 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

{vmercado, avillagra}@uaco.unpa.edu.ar, merreca@unsl.edu.ar

RESUMEN

El *análisis de autoría* (AA) es un área de investigación que ha ganado interés creciente en los últimos años principalmente por sus potenciales (y actuales) aplicaciones en problemas de seguridad nacional e inteligencia, lingüística forense, análisis de mercados e identificación de rasgos de personalidad, entre otros. El AA se enfoca en la clasificación automática de textos basándose fundamentalmente en las elecciones estilísticas de los autores de los documentos, e incluye distintas tareas de análisis como, por ejemplo: a) la *atribución de autoría*, b) la *verificación de autor*, c) la *detección de plagios*, d) la *determinación del perfil del autor* y e) la *detección de inconsistencias estilísticas*.

En este contexto, la línea de trabajo correspondiente a este trabajo de tesis de postgrado se propone el abordaje de dos tareas de AA, como lo son la *identificación del sesgo político* y *la atribución de autoría* en documentos periodísticos argentinos. Estas tareas de AA, serán complementadas con enfoques no-supervisados de aprendizaje automático como son el *agrupamiento de textos* (clustering) y la *modelización de tópicos*.

Palabras clave: Análisis Automático de Textos. Análisis de autoría. Determinación del perfil del autor. Identificación del sesgo político. Atribución de autoría. Análisis automático de documentos periodísticos.

CONTEXTO

Esta 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 de la Universidad Nacional de la Patagonia Austral, en el marco del Proyecto de Investigación 29/225 "Soluciones inteligentes para el desarrollo urbano sostenible" y particularmente en el marco del trabajo de tesis en la Maestría de Informática y Sistemas (MIS) de la Ing. Viviana Mercado.

1. INTRODUCCION

A partir de la disponibilidad de volúmenes inmensos de información en la Web, se reconoce cada día más el rol del *Aprendizaje Automático* y la *Extracción de Conocimiento a partir de los Datos* como herramientas fundamentales para hacer un uso adecuado y ventajoso de esta información. Esta tendencia crece día a día y se plantean nuevos escenarios relevantes como es el caso de *Big Data*, donde el contexto donde deben ser aplicados estos métodos es sumamente desafiante. En particular, un escenario que recibe creciente atención es el del *Análisis (automático) de Textos* (AAT, text analytics) o *aprendizaje automático con textos* (machine learning from text), lo cual resulta razonable si consideramos que gran parte de la información disponible en repositorios digitales, archivos de empresas y

organizaciones y la Web en general es de tipo textual.

El AAT surge de la interacción de 3 áreas principales: el Procesamiento del Lenguaje Natural, la Recuperación de la Información y el Aprendizaje Automático. El AAT incluye diversas tareas como el análisis de sentimiento/minería de opiniones, la extracción de información, la generación de resúmenes y el *análisis de autoría* (AA). El AA a su vez, comprende otras tareas más específicas como la *determinación del perfil del autor* (DPA) y la *atribución de autoría* (ATA). DPA es el área que identifica patrones compartidos por un grupo de gente y que aborda problemas de clasificación de acuerdo a la edad y género [Peersman et al., 2011, Schler et al., 2006, Argamon et al., 2009], nacionalidad, personalidad [Celli et al., 2014, Mairesse et al., 2007], orientación política [Abooraig et al.,2014, Conover et al.,2011, Malouf & Mullen, 2007], tendencias *suicidas* o *depresivas*, *rasgos de personalidad*, *perfiles de consumo* y *adicciones*, de *acosadores sexuales*, etc. La DPA, es un tema muy importante de investigación principalmente por sus potenciales (y actuales) aplicaciones en problemas de seguridad nacional e inteligencia, lingüística forense, análisis de mercados e identificación de rasgos de personalidad, entre otros. La ATA, por su parte, consiste en la atribución de un texto de autoría desconocida a uno de un conjunto de autores potenciales. Desde el punto de vista científico, esta tarea plantea varios desafíos interesantes ya que, a diferencia del análisis basado en los temas o tópicos de los textos, aquí se debe capturar información representativa de los *estilos de escritura* de los autores.

En este contexto, y como parte de un proyecto de tesis de postgrado, se ha propuesto el abordaje de dos tareas de AA, una de ATA y otra de DPA como lo son *la atribución de autoría* y la *identificación del sesgo político* en *documentos periodísticos argentinos*. Estas tareas de AA, serán complementadas con enfoques no-supervisados de aprendizaje automático como son el *agrupamiento* de textos (clustering) y la *modelización de tópicos*.

El resto de este artículo describe estas líneas de trabajo en la Sección 2 y los resultados esperados/obtenidos en la Sección 3.

2. LINEAS DE INVESTIGACION y DESARROLLO

En esta sección se describen las 3 líneas de investigación que se llevan a cabo como parte de esta tesis de postgrado, y que identificaremos de ahora en más como 1) *atribución de autoría en documentos periodísticos*, 2) *identificación del sesgo político en documentos periodísticos* y 3) *Análisis no supervisado de documentos periodísticos*.

2.1. *Atribución de autoría en documentos periodísticos*

La *atribución de autoría (ATA)* es una de las tareas principales dentro del *análisis de autoría* (AA) [Stamatatos, 2009]. El AA se enfoca en la clasificación automática de textos basándose fundamentalmente en las elecciones estilísticas de los autores de los documentos, e incluye distintas tareas de análisis como, por ejemplo: a) la *atribución de autoría*, b) la *verificación de autor*, c) la *detección de plagios*, d) la *determinación del perfil del autor* y e) la *detección de inconsistencias estilísticas*. Los enfoques predominantes en esta área están basados en el aprendizaje automático/de máquina supervisado. En pocas palabras, estos enfoques derivan, a partir de un conjunto de datos etiquetados (conjunto de entrenamiento) y un proceso inductivo de aprendizaje/entrenamiento, un clasificador que puede generalizar sus predicciones a otros datos no observados previamente. La representación clásica de los textos/documentos en estos casos, incluye tanto atributos basados en el contenido (palabras) como en el estilo de escritura de los autores.

Para el caso particular de llevar a cabo una tarea de ATA con documentos

periodísticos, el primer paso es la construcción de un "corpus" (colección de documentos) con los documentos de distintos periodistas. En nuestro caso, el foco fue en la recopilación de documentos de periodistas de reconocida adhesión a las políticas del gobierno Nacional en Argentina, en el período finalizado el 10 de Diciembre de 2015 y de documentos de periodistas que clara y abiertamente eran opositores a dichas políticas. Estos documentos, están originados en la información textual que dichos periodistas han hecho disponible en redes sociales, blogs, artículos en periódicos on-line, libros de investigación periodística, etc.

Así, la primera tarea fue identificar fuentes de información donde obtener textos periodísticos de Argentina, con una clara orientación política (oficialista vs opositor), como así también de los autores de los mismos. Posteriormente se procedió a recopilar toda esta información, compatibilizando los distintos formatos de los documentos (html, pdf, etc.) y generando un corpus con los textos "planos" de los mismos.

Para la ATA en documentos periodísticos analizaremos las particularidades que surgen para la identificación automática de autores, en aquellos contextos en donde los mismos tienen igual o diferente orientación política. En estos casos, se analizará cuáles son las "features" (estilográficas o de contenido) que son más relevantes para discriminar los distintos autores que pertenecen al mismo (o diferente) espectro político.

Como hipótesis de trabajo, se plantea que cuando la atribución de autoría se realice sobre la totalidad de los periodistas, tanto las características de contenido (palabras) como estilográficas serán relevantes. Sin embargo, cuando la atribución se restrinja a periodistas de la misma orientación política, el uso común de ciertas palabras y expresiones hará que los aspectos estilográficos tomen mayor relevancia para la identificación de los mismos.

2.2. *Identificación del sesgo político en documentos periodísticos*

La *identificación del sesgo político* (ISP) en un texto o documento es una forma de *determinación del perfil del autor* (DPA), otra de las tareas principales del análisis de autoría (AA). La ISP, al igual que otras tareas de DPA como la detección de personas depresivas o con diversos rasgos de personalidad, pedófilos y suicidas es una tarea desafiante dentro del análisis automático de textos ya que involucra, en general, el uso de representaciones de los textos que capturen aspectos de contenido y estilográficos de sus autores.

En este contexto, un área particular dentro de la ISP es la que se orienta al estudio de la orientación política en textos escritos por periodistas, y que referiremos de ahora en más como textos periodísticos (TP). Como ya se explicó previamente, consideraremos como textos periodísticos a aquella información que un periodista publica en diversos medios como puede ser un blog personal, un artículo escrito en un medio masivo como un diario o bien el contenido expresado en un libro de su autoría.

La ISP ha sido realizada con textos generados por usuarios comunes de medios sociales como Twitter [Cohen & Ruths, 2013], aunque más recientemente se ha realizado con los documentos producidos por periodistas [Lazaridou & Krestel, 2016]. Sin embargo, estos textos han sido escritos mayoritariamente en inglés o en otros idiomas, no existiendo, de acuerdo a nuestro conocimiento, estudios sobre ISP de documentos periodísticos en español.

En este trabajo realizaremos una primera aproximación a la ISP en textos periodísticos en español, en particular, de textos generados por periodistas argentinos como ya fue explicado en la sección previa. La tarea en este caso, será agrupar todos los documentos de periodistas "oficialistas" por un lado y "opositores" por el otro y convertir el problema de clasificación en uno de clasificación binaria ("oficialista" vs "opositores").

Nuestro objetivo aquí será responder las siguientes preguntas de investigación:

1) ¿Cuáles son las formas de representación de documentos más apropiadas para esta tarea?

2) ¿Cuáles son los algoritmos de aprendizaje más efectivos para usar con esas representaciones?

3) ¿Cuál es el impacto de los enfoques de reducción de dimensionalidad en las representaciones de los documentos?

4) ¿Cuánto se asemejan los resultados obtenidos con estudios similares con textos periodísticos escritos en otros idiomas?

2.2. *Análisis no supervisado de documentos periodísticos*

Las dos tareas de AA planteadas previamente son básicamente tareas de aprendizaje automático *supervisado,* es decir se basan en la información de *las clases* (periodistas en un caso, "oficialista" u "opositor" en el otro) que los documentos tienen asociadas. Sin embargo, existen otras formas de análisis automático de textos no supervisadas que suelen obtener información complementaria valiosa sobre las colecciones de documentos. Estos métodos, que incluyen el agrupamiento (*clustering*) no supervisado de documentos, la reducción no supervisada de dimensionalidad (por ejemplo, mediante *análisis de componentes principales*) o la *modelización de tópicos* se centran en analizar las propiedades de las "features" utilizadas para representar los documentos, sin hacer uso de la información de la clase asignada al documento.

El agrupamiento (clustering) de documentos, por ejemplo, busca generar *grupos* de documentos *relacionados*. La idea es que documentos dentro del mismo grupo sean similares (cercanos) entre sí (alta cohesión) y sean diferentes (alta separación) a los documentos de otros grupos.

Las técnicas de *reducción de dimensionalidad* como el *análisis de componentes principales* (PCA por sus siglas en inglés) permite transformar colecciones de datos de alta dimensionalidad a dos o tres dimensiones facilitando de esa manera su visualización.

Las técnicas de *modelización de tópicos* por su parte, como la *Latent Dirichlet Allocation* (LDA), intenta encontrar grupos de palabras (los *tópicos*) que aparecen en conjunto frecuentemente. Así, la LDA requiere que cada documento pueda ser entendido como una "mixtura" de un subconjunto de los tópicos.

En el contexto del análisis de documentos periodísticos, las tareas no supervisadas en este caso serán diversas:

1) En el contexto del agrupamiento, el objetivo será identificar cuáles son las representaciones y algoritmos de clustering que detectan en forma más efectiva los grupos oficialistas/opositores en un caso y los distintos periodistas en el otro.

2) Los métodos de reducción de dimensionalidad permitirán visualizar cuales son las clases que, a priori, lucen como más sencillas o complicadas para abordar por los enfoques supervisados. También se analizará el impacto de estas técnicas en la efectividad de los distintos clasificadores, considerando distintas dimensiones para los datos a analizar.

3) La modelización de tópicos por otra parte, permitirá identificar cuáles son los tópicos comunes a ambos tipos de periodistas (oficialistas y opositores) y cuales son propios de cada categoría.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Como resultado de esta tesis se espera lograr un *sistema integrado* de *atribución de autoría* con periodistas de la Argentina y de *determinación de la orientación política en documentos periodísticos*, que también soporte otras tareas de aprendizaje no supervisado como el *descubrimiento de tópicos* y el *agrupamiento* de documentos similares

Al tiempo presente, ya se ha generado un corpus de 200 documentos de 10 periodistas (5 oficialistas y 5 opositores) representándose los documentos con un enfoque de *bolsa de términos (bag of terms)*

variándose en cada caso el tipo de término utilizado (*n*-gramas de palabras, *n*-gramas de caracteres, términos "lematizados", términos "truncados", etc.). En cada caso se han utilizado distintos enfoques para el pesado de los términos como la ponderación *binaría*, ponderación *tf-idf*, etc. También se han representado los documentos utilizando todas las características del sistema *Language Inquiry Word Count* (LIWC). Los resultados preliminares en las tareas de atribución de autoría e identificación del sesgo político son muy prometedores y comparables a los obtenidos en tareas similares en otros idiomas como el inglés. En las tareas de análisis no supervisado ya se han obtenido 100 tópicos representativos de las temáticas abordadas por los distintos tipos de periodistas. En este momento se están comenzando los estudios con las restantes técnicas no supervisadas. Es importante observar, que si bien nuestro estudio está actualmente restringido al uso de técnicas "clásicas" en la representación de documentos, se prevé en un futuro considerar nuevos enfoques de *aprendizaje de representaciones*, que incluyan el aprendizaje de "*embeddings*" tales como **Word2vec**, **GloVe**, **fastText**, **StarSpace**, **ULM-FiT**, **ELMo** y **BERT**.

4. FORMACION DE RECURSOS HUMANOS

En esta línea de trabajo actualmente una integrante está desarrollando su tesis de Maestría correspondiente a la Maestría en Informática y Sistemas de la UNPA.

5. BIBLIOGRAFIA

Abooraig R., Alwajeeh A., Al- Ayyoub M., and Hmeidi I. **(2014)**. On the automatic categorization of arabic articles based on their political orientation. In *Proc. of the Third International Conference on Informatics Engineering and Information Science (ICIEIS2014)*.

Argamon S., Koppel M., Pennebaker J. W., and Schler J. **(2009)** *Automatically profiling the author of an anonymous text*. Communications of the ACM, 52:119123.

Celli F., Lepri B., Biel J.-I., Gatica- Perez D., Riccardi G., and Pianesi F.**(2014)**. *The workshop on computational personality recognition* 2014. In Proceedings of the ACM International Conference on Multimedia, MM '14, pages 1245-1246, New York, NY, USA. ACM.

Cohen R., and Ruths D. **(2013)** Classifying Political Orientation on Twitter: It's Not Easy!. Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media.

Conover M., Goncalves B., Ratkiewicz J., Flammini A., and Menczer F.**(2011)**. *Predicting the political alignment of twitter users*. In Proceedings of 3rd IEEE Conference on Social Computing (SocialCom).

Lazaridou K., and Krestel R. **(2016)** Identifying Political Bias in News Articles. TCDL Bulletin, vol. 12, 2016.

Mairesse F., Walker M. A., Mehl M. R. and Moore R. K. (**2007**). *Using Linguistic Cues for the Automatic Recognition of Personality in Conversation and Text*. In JAIR, 30, 457-500.

Malouf R. and Mullen T.**(2007)** *Graph-based user classification for informal online political discourse*.

Peersman C., Daelemans, W. and Van Vaerenbergh L. **(2011)**. *Predicting age and gender in online social networks.*In Proceedings of the 3rd international workshop on Search and mining user-generated contents, SMUC '11, pages 37-44, New York, NY, USA. ACM.

Schler J., Koppel M., Argamon S., and Pennebaker, J. W. **(2006)**. Effects of age and gender on blogging. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs,* pages 199-205, 2006.

Stamatatos E. (**2009**). *A Survey of Modern Authorship Atrribution Methods*. Journal of the American Society for Information science and technology. 60(3). 538-556.

Aprendizaje automático y visión por computadora

L. Lanzarini[1], C. Estrebou[1], F. Ronchetti[1,3], F. Quiroga[1,4], G. Camele[1,5], G. Rios[1,5], U. Cornejo Fandos[1,5], B. Rey[1,5], A. Rosete[6]

[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 postdoctoral UNLP [4] Becario postgrado UNLP [5] Pasante LIDI
[6] 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@lidi.info.unlp.edu.ar

CONTEXTO

Esta presentación corresponde a las tareas de investigación que se llevan a cabo en el III LIDI en el marco del proyecto "Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data" perteneciente al Programa de Incentivos (2018-2021) y del proyecto PITAP-BA "Computación de Alto Desempeño, Minería de Datos y Aplicaciones de Interés Social en la Provincia de Bs.As." evaluado y subsidiado por la Comisión de Investigaciones Científicas de la Provincia de Bs.As. (2017-2019).

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 clásicas, junto con las nuevas Redes Neuronales Convolucionales y Aprendizaje profundo. El trabajo presentado describe diferentes casos de aplicación en visión por computadora.

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.

En esta área se llevan a cabo diversas investigaciones. Por un lado, se están desarrollando dos librerías para potenciar la carga, el manejo y procesamiento de bases de datos de lengua de señas, así como de imágenes de formas de mano correspondientes a las mismas.

Por otro lado, se está estudiando cómo mejorar el reconocimiento de formas de manos de la lengua de señas con Redes Neuronales Convolucionales. Un problema importante dentro de esa área es el de otorgar invarianza a la rotación para reconocer formas de mano. A partir de ese requerimiento, se han desarrollado análisis sobre la forma en que las redes neuronales convolucionales aprenden la invarianza a la rotación y otras transformaciones afines, no sólo en el dominio de las formas de mano sino en general.

Siguiendo con la línea de reconocimiento de patrones en video, se realizó un estudio de diferentes modelos de clasificación de peatones en la vía pública. El objetivo de esta investigación se centró en el análisis de la generalización de los modelos, midiendo cómo se comportaban luego de ser entrenado con un conjunto de datos y ser evaluado con otro.

Por último, se está desarrollando un proyecto de I+D+I para identificar diferentes atributos faciales en rostros humanos como emociones y otras características.

Palabras clave: Aprendizaje Automático, Visión por Computadoras, Redes Neuronales, Lengua de Señas, Reconocimiento de Peatones.

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 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 procesar 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 [1,2].

El problema del reconocimiento de la lengua de seña es similar al del reconocimiento del habla, con la dificultad agregada de necesitar un procesamiento multimodal y la falta de bases de datos orientadas al entrenamiento de un sistema de reconocimiento. Además, las bases de datos disponibles varían ampliamente en el formato, metadatos, anotaciones y forma de obtención. Por estos motivos, el campo de la lengua de señas no ha logrado desarrollar sistemas con un nivel de desempeño similar al alcanzado por el reconocimiento del habla.

En esta área, en donde se ha terminado una tesis doctoral y otra está en progreso, se está trabajando en crear una librería para unificar y simplificar la carga y el preprocesamiento y etiquetado de bases de datos de lengua de seña dinámicas[1]. Se proveerá soporte para cargar una base de datos de Lengua de Señas Argentina desarrollada previamente [2], y otras más.

La librería a desarrollar buscar proveer una interfaz unificada para cargar y unir bases de datos, de modo de poder ampliar el repertorio de señas y de ese modo poder entrenar modelos más potentes. Al mismo tiempo, facilitará el ingreso a la investigación en el área, incrementando el nivel y la cantidad de los desarrollos.

1.2. Clasificación de formas de mano de la Lengua de Seña

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 [1,3], 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.

Las bases de datos de formas de manos para lengua de señas presentan problemas similares a las bases de datos completas de lengua de señas. Por ende, también se está trabajando en una librería para unificar su carga y preprocesamiento[2]. Esto es más importante aún de cierto modo para las formas de mano, porque si bien las lenguas de señas son regionales y un modelo entrenado para un país o región no sirve para otro, el conjunto de formas de mano es universal para todos los seres humanos, si bien cada lengua utiliza subconjuntos distintos.

Por otro lado, las Redes Neuronales Convolucionales (CNN, por sus siglas en

[1] https://github.com/midusi/sign_language_datasets

[2] https://github.com/midusi/handshape_datasets/

inglés) son actualmente el estado del arte en el área. En un estudio previo, se realizaron experimentos comparando diversas arquitecturas de CNN para clasificar formas de manos, pero sólo con dos bases de datos, LSA16 y RWTH [4]. Actualmente, se está extendiendo este análisis a varias bases de datos mediante el uso de la librería mencionada anteriormente.

1.3. Invarianza a la rotación en Redes Convolucionales

Una de las propiedades deseables en un modelo de clasificación de formas de mano es que sea invariante a la rotación, debido a que la forma de mano se considera la misma independientemente de la orientación que tenga.

Esto motivó el estudio de distintos modelos y técnicas que permitan que una CNN sea invariante a la rotación y otras transformaciones afines. En los últimos años, varios modelos han sido propuestos para añadir invarianza a la rotación en CNNs [6].

No obstante, no existe una comparación sistemática de dichos modelos, en especial contra un método simple y efectivo como la generación de datos transformados (*data augmentation)*. Por ende, se realizó una comparación de los modelos más emblemáticos del área con y contra *data augmentation*. Para la comparación, se utilizaron dos modelos de CNN: una red convolucional simple, y otra completamente convolucional. Los modelos fueron entrenados y probados con MNIST y CIFAR10, las dos bases de datos de clasificación más conocidas, con el objetivo de que los resultados sean fáciles de interpretar. Encontramos que estos modelos no ofrecen un desempeño significativamente superior vs *data augmentation,* aunque si pueden reducir el número de parámetros del mismo. Además, se analizó la posibilidad de aplicar modelos entrenados sin invarianza y reentrenarlos para que obtengan dicha invarianza. Se encontró que, sorpresivamente, las capas convolucionales son capaces de reentrenarse para ofrecer invarianza, y no solo las capas densas o completamente conexas pueden hacerlo, no obstante el hecho de que individualmente una capa convolucional no puede ser nunca invariante a la rotación[6].

Actualmente, se está ampliando la investigación en esta área, desarrollando un método para cuantificar la (in)varianza de cada unidad en una CNN de forma de comprender mejor como se codifica la invarianza en las redes convolucionales.

1.4. Detección de peatones

En el marco de la investigación sobre detección de objetos en video, se realizó un proyecto sobre detección de peatones en la vía pública. El objetivo estuvo centrado en el estudio de la transferencia de aprendizaje de los modelos entrenados para validar su efectividad al utilizarlos en entornos reales. Se eligieron tres bases de datos para entrenar los modelos de clasificación: INRIA, DAIMLER y TUD-Brussels, sumando más de 5.000 imágenes en total.

Luego, se analizaron diferentes descriptores con el fin de determinar su utilidad para incrementar la capacidad predictiva del modelo. Se estudiaron los Histograma de Gradientes Orientados (HOG) y Patrones Binarios Locales (LBP), siendo los más utilizados para estas tareas en el estado del arte. En tercer lugar, se entrenó una Máquina de Vectores de Soporte (SVM) con el fin de clasificar las imágenes de entrenamiento.

Se estableció un protocolo de experimentación para entrenar un modelo con cada conjunto de datos y evaluarlo en todos los restantes. Del mismo modo, se realizó un entrenamiento con diferentes conjuntos de datos mezclados para verificar si esto enriquecía la generalización del modelo entrenado.

Los resultados obtenidos mostraron que, si bien cada conjunto de datos presentaba escenas del mundo real, existen diferencias significativas que hacen que un modelo entrenado con un conjunto de imágenes no funcione apropiadamente con otro. Por otro lado, se encontró que al entrenar un modelo

con la combinación de diferentes bases de datos se genera una mayor transferencia de aprendizaje, aunque no siempre ayuda al entrenamiento de un conjunto de datos particular. Los resultados de esta investigación se encuentran publicados en [7].

1.5. Clasificación de atributos faciales

En el marco del proyecto de investigación, desarrollo e innovación "Análisis de Atributos Faciales" evaluado y financiado por la Facultad de Informática de la UNLP y continuando la línea de trabajos anteriores de reconocimiento y localización de objetos en video, se está trabajando en el desarrollo de aplicaciones capaces de reconocer atributos faciales en rostros humanos, como pueden ser emociones u otras características.

Para llevar esto a cabo, se están analizando diferentes bases de datos utilizadas en el estado del arte, como Facial Expressions in the Wild que contiene más de 80 mil imágenes obtenidas de internet con etiquetas de emociones básicas como *alegría, tristeza* o *enojo.* Otro conjunto de datos analizado es Large-scale CelebFaces Attributes (CelebA) Dataset, que contiene más de 200 mil imágenes de personalidades reconocidas con diferentes etiquetas sobre el rostro como el color del cabello, si tiene bigote, si usa anteojos, etc.

Como modelos de clasificación se están analizando diferentes arquitecturas de CNNs actuales, como VGG, Inception o ResNet. Si bien el grupo de trabajo ya ha realizado trabajos con estas arquitecturas, aún resta aplicarlas al procesamiento en tiempo real de rostros humanos.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio de técnicas para obtener invarianza a transformaciones afines en clasificadores de imágenes.
- Representación y clasificación de configuraciones de manos para el lenguaje de señas.
- Representación y clasificación de señas dinámicas.
- Estudio y desarrollo de modelos de Redes Convolucionales Profundas aplicados a problemas de visión por computadora.
- Estudio de la generalización de los modelos de detección de peatones.
- Análisis de diferentes representaciones aplicados a la clasificación de emociones y atributos faciales.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Desarrollo de dos librerías para potenciar la investigación en reconocimiento de lengua de señas y clasificación de formas de mano.
- Comparación de varios modelos de redes convolucionales con invarianza a las transformaciones afines.
- Análisis de los mecanismos de las redes neuronales para adquirir invarianza a las transformaciones afines.
- Modelos de clasificación de peatones y sistema de detección en tiempo real.
- Sistema de clasificación de emociones humanas en rostros y descripción de atributos faciales.

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, 3 tesistas y 1 profesor extranjero.

Dentro de los temas involucrados en esta línea de investigación, en los últimos dos años se han finalizado 2 tesis de doctorado, 1 tesis de especialización, y 7 tesinas de grado de Licenciatura.

Actualmente se están desarrollando 1 tesis de doctorado, 1 tesis de especialista y 5 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., Estrebou C., Lanzarini L., Rosete A. *Sign languague 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., Estrebou, C.A., Lanzarini. *Handshape recognition for Argentinian Sign Language using ProbSom*. Journal of Computer Science & Technology, vol. 16, Nº 1, págs. 1-5, ISSN 1666-6038, 2016.

[3] Ronchetti, F., Quiroga, F., Estrebou, 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] 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).

[5] 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).

[6] Quiroga F., Ronchetti F., Lanzarini L., Fernandez-Bariviera A. *Revisiting Data Augmentation for Rotational Invariance in Convolutional Neural Networks*. International Conference on Modeling and Simulation in Engineering, Economics and Management (MS'2018 GIRONA) (en prensa).

[7] Camele G, Quiroga F., Ronchetti F., Hasperué W., Lanzarini L. *Transferencia de Aprendizaje para Clasificación de Peatones.* XXIV Congreso Argentino de Ciencias de la Computación. CACIC 2018. Tandil. Octubre 2018.

Desarrollo de algoritmos genéticos utilizando diferentes frameworks de MapReduce: MPI vs. Hadoop

Juan José Barbero, Martín Tamagusku, Hugo Alfonso[1],
Carlos Bermudez, Gabriela Minetti[1], Carolina Salto[1]
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 MapReduce es un paradigma popular, que permite a los usuarios no especializados utilizar grandes plataformas computacionales paralelas de manera transparente. Hadoop es la implementación más utilizada de este paradigma y, de hecho, para una gran cantidad de usuarios, la palabra Hadoop y MapReduce son intercambiables. Pero, hay otros framewoks que implementan este paradigma de programación, como MapReduce-MPI. Dado que las técnicas de optimización pueden beneficiarse enormemente de este tipo de modelado informático de uso intensivo de datos, en esta línea de investigacón analizamos el efecto del rendimiento del desarrollo de algoritmos genéticos (GA) utilizando diferentes marcos de MapReduce (MRGA).

Palabras claves: Metaheurísticas, optimización, operadores

Contexto

Esta línea de investigación se desarrolla en el marco del 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. En particular, la línea de investigación aquí descripta se lleva adelante conjuntamente con los Dres. Alba y Luque de la Universidad de Málaga.

Introducción

MapReduce (MR) es un paradigma de programación, creado por los investigadores de Google [1], para procesar y generar grandes conjuntos de datos con un algoritmo paralelo y distribuido en clusters y en la nube. El programador puede abstraerse de los problemas de la programación distribuida y paralela, ya que MR administra el equilibrio de carga, el rendimiento de la red y la tolerancia a fallas en algunos frameworks. Esto hizo que MR fuera popular para aquellos que no sabían sobre programación paralela pero que querían usarla, creando una nueva rama del paralelismo donde el enfoque está en la aplicación y no en la explotación del hardware subyacente.

La implementación de MR en Google es una librería propietaria en C++, con comunicación entre máquinas en red a través de llamadas a procedimientos remotos. Permite la tolerancia a fallas cuando se emplean grandes cantidades de máquinas, y puede usar discos como memoria fuera del núcleo para procesar conjuntos de datos a escala de petabytes [2]. HADOOP incorpora también estas características en su implementación [3], pero es una librería de código abierto en Java

que, además, soporta varios lenguajes de programación.

Puesto que que Hadoop permite el paralelismo de datos y control, es una ciencia común preguntarnos sobre otras herramientas de software que realicen trabajos similares. MapReduce-MPI (MR-MPI) [4] es una librería construida sobre MPI con un objetivo similar. Al tratarse de una librería de código abierto en C++, permite un control preciso sobre la memoria y el formato de los datos asignados por cada procesador durante una operación de MapReduce. Tanto *map* como *reduce* son funciones especificadas por el usuario que se pueden escribir en los lenguajes C, C++ y Python. Aquí se puede tener más control sobre la plataforma, lo que permite mejorar el rendimiento del ancho de banda y reducir los costos de latencia. Si bien sus servicios (para Hadoop y MR-MPI) no son exactamente iguales, la mayoría de las aplicaciones se pueden escribir sobre ellas, y se hace un análisis comparativo para cualquier científico curioso, a fin de ofrecer evidencias sobre su desempeño.

Dada la necesidad científica e industrial de nuevos algoritmos evolutivos, en particular de los algoritmos genéticos (GA) escalables, un modelo informático de uso intensivo de datos puede beneficiar enormemente el campo para construir nuevos modelos de optimización y aprendizaje automático con ellos. Por esta razón, el objetivo de esta línea de investigación es analizar diferentes formas de introducir los elementos de MR en un algoritmo genético simple (SGA) [5]. Los primeros resultados de este trabajo se cristalizan en una técnica llamada MRGA [6], que surge de una paralelización de cada iteración de un SGA bajo el Modelo Paralelo de Nivel de Iteración [7].

Desarrollo

En esta sección se describe la primera parte del estudio propuesto en esta línea de investigación. Para ello es necesario realizar una breve descripción del paradigma MR.

El paradigma MapReduce organiza una aplicación en un par (o una secuencia de pares) de las funciones *map* y *reduce* [1]. Las funciones *map* son independientes entre sí y se ejecutan en paralelo sin conexión entre ellas. Las segundas unen los valores con la misma clave.

Cada función *map* toma como entrada un conjunto de pares clave/valor (registros) de los archivos de datos y genera un conjunto de pares clave/valor intermedio. Luego, MR agrupa todos estos valores intermedios asociados con una misma clave. Un grupo de valores y su clave asociada es la entrada a la función *reduce*, que combina estos valores para producir un conjunto nuevo y posiblemente más pequeño de pares clave/valor que se guardan en archivos de datos. Además, esta función recibe los valores intermedios a través de un iterador, lo que permite al modelo manejar listas de valores que son demasiado grandes para formar parte de la memoria principal.

Los datos de entrada se dividen automáticamente en un conjunto de M divisiones cuando las invocaciones *map* se distribuyen en múltiples máquinas, donde cada división de entrada se procesa mediante una invocación *map*. El espacio de claves intermedias se divide en R piezas, por medio de la función de partición (por ejemplo, $hash(key)\mathbf{mod}R$) que se distribuyen en R invocaciones de *reduce*.

MRGA se basa en el SGA propuesto por [5], que solo utiliza los operadores de selección y recombinación. El algoritmo propuesto preserva el comportamiento de SGA, pero recurre a la paralelización de algunas partes, tales como: la evaluación y la aplicación de operadores genéticos. Si bien, MRGA realiza varias operaciones en paralelo, su comportamiento es el mismo que el de un GA secuencial.

Se utiliza un par clave/valor para representar a individuos en la población. La representación de un individuo para el problema de la mochila es una secuencia de bits, una para cada elemento, que indica si el elemento está en la mochila o no. A continuación, comienza el proceso evolutivo iterativo, pero el modelo MR no ofrece métodos para implementar algoritmos iterativos. En consecuencia, se genera una cadena de múltiples tareas MapReduce utilizando la salida de la última como la entrada de la siguiente. Una consideración importante es que la condición de terminación debe computarse dentro del proceso principal del programa MapReduce. Cada iteración consta de las funciones *map* y *reduce*.

Teniendo en cuenta los problemas de gran tamaño, los individuos grandes y las funciones de evaluación complejas, las funciones *map* calculan la aptitud de los individuos. Como a cada *map*, MR le ha asignado una porción diferente de datos, se evalúa un conjunto de diferentes individuos en paralelo. Esta aptitud se agrega como valor en el par clave/valor.

Las funciones *reduce* llevan a cabo las operaciones genéticas. La selección del torneo binario se realiza localmente con el espacio clave intermedio, que se distribuye en la etapa de partición después de la operación *map*. El operador de cruce uniforme (UX) se aplica sobre los individuos seleccionados. El reemplazo generacional se implementa al escribir estos nuevos individuos en un archivo para convertirse en la entrada del *map* en la próxima tarea MapReduce (una nueva iteración).

Resultados obtenidos

En este primera etapa, se estudian diferentes formas de introducir los elementos MapReduce en un algoritmo genético simple. Como consecuencia, surgieron dos implementaciones de MRGA utilizando ambos frameworks de código abierto MapReduce, Hadoop y MR-MPI, para generar los algoritmos MRGA-H y MRGA-M respectivamente. Estos algoritmos resuelven un problema de gran tamaño de interés industrial como es el problema de la mochila. Cuando se analiza la eficiencia, detectamos que los dos algoritmos MRGA necesitan el mismo número de iteraciones para encontrar sus mejores soluciones. La calidad de las soluciones obtenida por los algoritmos es bastante similar, observándose una pequeña ventaja a favor de MRGA-M. Este es un resultado esperado porque ambos algoritmos MRGA se basan en el mismo algoritmo genético selecto-recombinativo.

Considerando la escalabilidad algorítmica desde el punto de vista de la dimensión del problema, observamos que MRGA-H es más escalable que MRGA-M. Las razón de esto es que el crecimiento del tiempo requerido por MRGA-H no es directamente proporcional al incremento en la dimensión de la instancia, es decir, el porcentaje de crecimiento del tiempo de ejecución disminuye a medida que aumenta el número de elementos.

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 tesinas 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] J. Dean and S. Ghemawat, "MapReduce: simplified data processing on large clusters," in *OSDI'04: Proceedings of the 6TH Conference on Symposium on Operating Systems Desing and Implementation*. USENIX Association, 2004.

[2] ——, "MapReduce: Simplified data processing on large clusters," *Commun. ACM*, vol. 51, no. 1, pp. 107–113, Jan. 2008. [Online]. Available: http://doi.acm.org/10.1145/1327452.1327492

[3] "Welcome to Apache™ Hadoop®!" The Apache Software Foundation, http://hadoop.apache.org/, Tech. Rep., 2014.

[4] S. Plimpton and K. Devine, "Mapreduce in MPI for large-scale graph algorithms." *Parallel Computing*, vol. 37, no. 9, pp. 610–632, 2011.

[5] D. Goldberg, *The Design of Innovation: Lessons from and for Competent Genetic Algorithms*. Kluwer Academic Publishers, 2002.

[6] C. Salto, G. F. Minetti, E. Alba, and G. Luque, "Developing genetic algorithms using different mapreduce frameworks: MPI vs. hadoop," in *CAEPIA*, ser. Lecture Notes in Computer Science, vol. 11160. Springer, 2018, pp. 262–272.

[7] E. Talbi, *Metaheuristics: From Design to Implementation*. Wiley Publishing, 2009.

DESARROLLO DE HERRAMIENTAS PARA LA OPERABILIDAD DE PROCESOS

Enrique E. Tarifa, Sergio L. Martínez, Samuel Franco Domínguez, Susana A. Chalabe, Luis E. Ituarte, Álvaro F. Núñez, Cristian D. Yurquina, Ubaldo J. M. Aramayo, Flavia A. Copa

Facultad de Ingeniería / Universidad Nacional de Jujuy / CONICET
Ítalo Palanca Nº 10 / S. S. de Jujuy / Provincia de Jujuy / Tel. 0388-4221591
eetarifa@fi.unju.edu.ar; smartinez@fi.unju.edu.ar; sfdominguez@yahoo.com.ar; susana.chalabe@gmail.com; luisituarte@hotmail.com; afnunez@fi.unju.edu.ar; c.d.yurquina@gmail.com; uba.aramayo@gmail.com; ayelencopa20@gmail.com

Área temática: Agentes y Sistemas Inteligentes (ASI)

RESUMEN

El proyecto *"Desarrollo de Herramientas para la Operabilidad de Procesos"* tiene como objetivo general desarrollar herramientas para la operación de procesos, enfocado principalmente hacia los procesos químicos e industriales. Para ello, se diseñarán, adaptarán y/o aplicarán herramientas propias de la Operabilidad de Procesos (flexibilidad, controlabilidad, confia-bilidad, robustez). Los procesos que tendrán prioridad son los que están implementados en la planta piloto de la Facultad de Ingeniería de la UNSa (extracción líquida-líquida, absorción gas-líquida, producción de vapor, pasteurización, reacción, entre otros), y los procesos vinculados al gas, al petróleo, al litio y a las energías no convencionales —especialmente, a la energía solar—.

Los objetivos particulares del proyecto propuesto son los siguientes: 1) Desarrollar métodos para determinar y aumentar la flexibilidad de procesos; 2) Desarrollar sistemas de supervisión; 3) Desarrollar sistemas de control avanzados; 4) Desarrollar material para la enseñanza de Ingeniería Química (simula-dores, optimizadores, sistemas de control, publicaciones y apuntes). Para alcanzar los objetivos propuestos, se emplearán técnicas del campo de la Ingeniería de Procesos (simulación, optimización, diseño) y de la Inteligencia Artificial (sistemas expertos fuzzy, redes neuronales, minería de datos).

Palabras clave: Procesos industriales, Operabilidad, Flexibilidad, Controlabilidad, Confiabilidad.

CONTEXTO

El grupo responsable de la realización del proyecto de referencia, es el grupo de investigación *IngProAr* (Ingeniería de Procesos Argentina), con lugar de trabajo en la Facultad de Ingeniería de la UNJu (Universidad Nacional de Jujuy). Este grupo fue creado en 1995 por el Dr. Enrique Tarifa (investigador de CONICET y de la UNJu), y está constituido por docentes investigadores, profesionales y estudiantes de diversas áreas disciplinares. Desde sus inicios, este grupo trabajó en estrecha colaboración con el instituto de desarrollo y diseño INGAR (CONICET, Santa Fe), el instituto de catálisis INCAPE (CONICET, Santa Fe) y el grupo de petroquímica de la Facultad de Ingeniería de la UNSa.

El proyecto al que se hace referencia es continuación del proyecto *"Desarrollo de Herramientas para la Operabilidad de Procesos Productivos"*, que fue financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales (SeCTER) y la Facultad de Ingeniería de la UNJu, y fue reconocido por el Programa de Incentivos de la Secretaría de Políticas Universitarias (SPU). El nuevo proyecto, denominado *"Desarrollo de Herramientas para la Operabilidad de Procesos"*, también es financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales (SeCTER) y la Facultad de Ingeniería, ambas de la Universidad Nacional de Jujuy, y fue reconocido por la SeCTER-UNJu con el código D/0157 y por el Programa de Incentivos con el código 08/D157.

En este nuevo proyecto, se plantea también la cooperación con el recientemente creado Centro de Desarrollo Tecnológico "General Manuel Savio" (UNJu, Gobierno de la Provincia de Jujuy, y el Ministerio de Ciencia, Tecnología e Innovación productiva de la Nación).

1. INTRODUCCIÓN

El objeto de estudio del proyecto propuesto es la operabilidad de procesos químicos e industriales. La operabilidad de un proceso, especialmente referido a procesos industriales o tecnológicos, comprende las siguientes propiedades:

- **Flexibilidad:** Representa la capacidad del diseño de un sistema para obtener una operación en estado estacionario estable en un rango de condiciones inciertas que pueden encontrarse durante la operación de la planta, y es el primer paso que se debe considerar para la operabilidad de un diseño.
- **Controlabilidad:** Puede ser definida como la facilidad con la cual un proceso continuo puede ser mantenido en un estado estable específico o deseado.
- **Confiabilidad:** Hace referencia a la capacidad de todos los elementos que conforman la planta para funcionar continuamente por un periodo de tiempo según un conjunto de especificaciones o condiciones.
- **Robustez:** Es la capacidad de un sistema de hacer mínima la variación de la medida de calidad de los productos ante variaciones de las condiciones de operación.

Debido a que las propiedades enunciadas son importantes por sí mismas, existen trabajos que se centran en cada una de ellas. Los estudios de flexibilidad suelen enfocarse en la definición y posterior determinación de índices que capturen apropiadamente esta característica del proceso. Entre las acciones a tomar para aumentar la flexibilidad de un proceso está la modificación de su diseño. En el aspecto de la controlabilidad, los estudios se orientan al planteo y verificación de condiciones que aseguren tal condición de los procesos. Entre las acciones a tomar para mejorar la controlabilidad del proceso está la modificación del diseño y la implementación de un adecuado sistema de control. En el aspecto de la confiabilidad, los estudios tienen como objetivo determinar la probabilidad de ocurrencia de fallas y sus consecuencias. En este caso, las acciones a tomar son preventivas (destinadas a disminuir la probabilidad de falla y sus consecuencias) y paliativas (producida la falla, se toman las medidas necesarias para atenuar sus efectos). En el aspecto de la robustez, los trabajos se concentran en el diseño de sistemas de control apropiados que se comporten adecuadamente en un rango amplio de condiciones de operación. Finalmente, debido a que las acciones que se tomen para mejorar una propiedad también afectan a las restantes, existen trabajos que estudian simultáneamente dos o más de las citadas propiedades.

Es de destacar que los actuales estudios de operabilidad involucran cada vez más el uso de técnicas de Inteligencia Artificial (redes neuronales, sistemas expertos, lógica *fuzzy*, entre otras), siendo los sistemas de control inteligentes los mejores exponentes de esta situación. También, se han aplicado con éxito estas técnicas para mejorar el diseño de los procesos, recurriendo en especial a estrategias evolutivas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto que se describe se enmarca en las siguientes líneas prioritarias de la UNJu (Res. CS N°168/93):

Desarrollo Económico–Social Regional Sustentable: Las herramientas a desarrollar en el proyecto servirán para aumentar la operabilidad de los procesos industriales y tecnológicos de la región. En consecuencia, se favorecerá el desarrollo económico de la región.

La Facultad de Ingeniería de la UNJu, también ha establecido líneas prioritarias de investigación, donde este proyecto tiene un importante sustento (Res. CAFI N°393/13):

Línea 1 - La exploración, manejo, valorización, obtención, transformación y aprovechamiento de recursos naturales renovables, no renovables y energéticos. La Gestión ambiental relacionada: La vinculación con esta línea se presenta debido a que las herramientas a desarrollar en el presente proyecto pueden ser aplicadas a cualquier tipo de proceso productivo o tecnológico. Además, entre los procesos a estudiar, están los

que involucran petróleo, gas, energía solar y litio; los cuales son los recursos energéticos más ampliamente utilizados en la actualidad.

Línea 3 - El estudio de procesos específicos o integrados que contribuyan a la cadena de valor de los productos obtenidos: Esta línea comprende los siguientes temas: Análisis, diseño y síntesis óptima de procesos; simulación y optimización de procesos; operación y mantenimiento de procesos; supervisión y control de procesos; sistemas de apoyo a la toma de decisiones en procesos; integración de masa y energía; análisis del comportamiento de procesos. Los resultados a obtener con el proyecto propuesto están vinculados a la mayoría de los temas citados.

Línea 5 - El desarrollo de la Informática, los sistemas y la tecnología derivados de ésta: Algunos de los productos del proceso son aplicaciones informáticas —simuladores, optimizadores, sistemas de control—, para las cuales se deben desarrollar nuevos algoritmos numéricos o técnicas de Inteligencia Artificial.

Línea 6 - El desarrollo de la calidad educativa y el estudio de los procesos de enseñanza y aprendizaje: Algunos productos del proyecto —simuladores, optimizadores, sistemas de control, publicaciones, apuntes— tienen impacto directo sobre la calidad educativa y serán empleados en el dictado de clases, desarrollo de trabajos prácticos, seminarios o conferencias.

3. OBJETIVOS Y RESULTADOS

El proyecto en cuestión, tiene como objetivo general desarrollar herramientas para la operación de procesos. Para ello se diseñarán, desarrollarán, adaptarán y aplicarán herramientas propias de la Operabilidad de Procesos.

Los objetivos particulares del proyecto propuesto son los siguientes: 1) Desarrollar métodos para determinar y aumentar la flexibilidad de procesos; 2) Desarrollar sistemas de supervisión; 3) Desarrollar sistemas de control avanzados; 4) Desarrollar material para la enseñanza de Ingeniería Química (simuladores, optimizadores, sistemas de control, publicaciones y apuntes). Para alcanzar los objetivos propuestos, se emplearán técnicas del campo de la Ingeniería de Procesos (simulación, optimización, diseño) y de la Inteligencia Artificial (sistemas expertos fuzzy, redes neuronales, minería de datos).

Para el nivel de desarrollo actual del proyecto, se realizaron diversas actividades como se detalla a continuación.

Publicaciones en revistas científicas

- Ituarte, L. E., *"*Componentes Electrónicos en Plantas Fotovoltaicas: El Caso del Inversor Multi-Cascada*"*, *Proyección*, N° 83. Marzo 2017, Colegio de Ingenieros de Jujuy, 2018.
- Busto M., Tarifa E. E., Vera C. R., *"*Coupling Solvent Extraction Units to Cyclic Adsorption Units*"*, *International Journal of Chemical Engineering*, Volume 2018 (2018), Article ID 1620218, 17 pages, https://doi.org/10.1155/2018/1620218
- Tarifa E. E., Franco Domínguez S., Martínez S. L., *"*Determinación de la instrumentación óptima para un sistema*"*, *Difusiones*, N° 14, 45-62, ISSN: 2314-1662, 2018.
- Ituarte L. E., Martínez S. L., Tarifa E. E., *"*Heliostatos y el control automático de orientación: una opción de mejora para las plantas de energía solar de la provincia de Jujuy*"*, *Difusiones*, N° 14, 63-70, ISSN: 2314-1662, 2018.
- Gutiérrez J. P., Tarifa E. E., Erdmann E., *"*Steady-state energy optimization and transition assessment in a process of CO_2 absorption from natural gas*"*, *Energy*, Vol. 159, 15 September 2018, 1016-1023, https://doi.org/10.1016/j.energy.2018.06.108, 2018.
- Tarifa E. E., Martínez S. L., Franco S., Argañaraz J. F. *"*Formulation of an optimal academic exam*"*, *Journal of Computer Science & Technology*, vol. 18, no. 2, pp. 167-177, DOI: 10.24215/16666038.18.e19, 2018.

Trabajos en Congresos

- Tarifa E. E., Núñez A., Martínez S., *"*Análisis de sensibilidad en un evaporador para jugo de tomate*"*, *IV Jornadas de Jóvenes Investigadores*, 25 y 26 de octubre de 2018, Facultad de Humanidades y Ciencias Sociales de la UNJu, San Salvador de Jujuy, 2018.
- Benitez L. A., Erdmann E., Tarifa E. E., *"*Desarrollo de Procedimientos Óptimos de Operación para el Sector de Ajuste del Punto de Rocío en Plantas de Acondicionamiento de

Gas Natural*"*, *IV Jornadas de Jóvenes Investigadores*, 25 y 26 de octubre de 2018, Facultad de Humanidades y Ciencias Sociales de la UNJu, San Salvador de Jujuy, 2018.

- Aramayo U. J., Martínez S. L., Tarifa E. E., *"*Modelos de redes neuronales utilizados para predicción*"*, *IV Jornadas de Jóvenes Investigadores*, 25 y 26 de octubre de 2018, Facultad de Humanidades y Ciencias Sociales de la UNJu, San Salvador de Jujuy, 2018.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo que forma el grupo de investigación *IngProAr* cuenta con profesores, auxiliares, graduados y pasantes de la Facultad de Ingeniería, quienes desarrollan diversas actividades en la formación de recursos humanos, durante la ejecución del proyecto. Particularmente, los docentes investigadores del grupo *IngProAr* están a cargo de diversas cátedras en distintas carreras de la Facultad de Ingeniería de la UNJu: *"Simulación y Optimización de Procesos"*, *"Ingeniería de Procesos"*, *"Electricidad y Electrónica"* e *"Introducción a la Informática"* —en Ingeniería Química—; *"Métodos de Simulación"*, *"Inteligencia Artificial"*, *"Técnicas y Estructuras Digitales"*, *"Metodología de la Programación"* —en Ingeniería Informática—; *"Electrotecnia"* en Ingeniería Industrial; *"Geología Ambiental"* y *"Cartografía especial"* en Ingeniería de Minas. Es de destacar que los contenidos de dichas materias están directamente vinculados con los temas considerados en este proyecto. Como actividades de formación de recursos humanos, se pueden destacar:

Tesis de posgrado

- Tarifa E. E., codirección de la tesis doctoral *"Secado en lecho de chorro bidimensional para la deshidratación de proteínas del plasma y porción globina de sangre bovina"*, Ing. Lara Valeria Lescano Farías, Directora: Dra. Eve Liz Coronel. Res. F.A.A. N°641/2013, UNSE, Santiago del Estero.
- Tarifa E. E., dirección de la tesis de maestría *"Diseño de un Sistema de Gestión Ambiental para la empresa de tratamiento de líquidos cloacales El Cadillal"*, Ing. Karina Alejandra Palma, Res. RI N°97/15, UCASAL, Salta.
- Tarifa E. E., dirección de la tesis doctoral *"Control Inteligente con Algoritmos Híbridos Optimizados aplicados a Modelos de Procesos Productivos"*, Ing. Sergio L. Martínez, Codirector: Dr. Juan P. Gruer (UNT), Facultad de Ciencias Exactas y Tecnología, Universidad Nacional de Tucumán, Res. FACET 620/2017, desde may/2017.
- Tarifa E. E., dirección de la tesis doctoral *"Desarrollo de procedimientos óptimos de operación para el sector de ajuste de punto del rocío en plantas de acondicionamiento de gas natural"*, Ing. Leonel Alberto Benítez, UNSa, beca doctoral de CONICET, tesis aprobada el 21 de mayo de 2018.

Tesis de grado

- Martínez S. L., dirección y codirección de la tesis de grado *"Aplicación de Redes Neuronales Artificiales: Tratamiento de series temporales estacionales para predicción de alturas y caudales en Caimancito, Jujuy, Argentina."*, Tesista: Ubaldo Juan Manuel Aramayo, Carrera de Ingeniería Informática, Facultad de Ingeniería, UNJu. Res. FI N°052/17. En ejecución.
- Martínez S. L., Dirección de la tesis de grado *"Desarrollo de una Herramienta con Interfaz Gráfica para Implementación de Redes Neuronales Feedforward"*. Tesista: David Llusco, codirector Ing. Jorge J. Gutiérrez, Carrera de Ingeniería Informática, UCSE-DASS, San Salvador de Jujuy. En ejecución.

Becarios de Investigación

- Cristian D. Yurquina: Estudiante de Ingeniería Química en la Facultad de Ingeniería de la UNJu, Estudiante Colaborador de *IngProAr*.
- Ubaldo J. M. Aramayo: Tesista y Estudiante de Ingeniería Informática en la Facultad de Ingeniería de la UNJu, Estudiante Colaborador de *IngProAr*.
- Flavia A. Copa: Estudiante de Ingeniería Informática en la Facultad de Ingeniería de la UNJu, Alumno Colaborador de *IngProAr*.

5. BIBLIOGRAFÍA

Azar A., Vaidyanathan S., *Handbook of Research on Advanced Intelligent Control Engineering and Automation*, Editorial Advisory Board, ISBN13: 9781466672482, 2015.

Banaja Mohanty, Sidhartha Panda, P.K. Hota, "Controller parameters tuning of differential evolution algorithm and its application to load frequency control of multi-source power system", *International Journal of Electrical Power & Energy Systems*, Volume 54, Pages 77-85, 2014.

Bidkhori H., Simchi-Levi D., Wei Y., "Analyzing process flexibility: A distribution-free approach with partial expectations", *Operations Research Letters*, vol. 44, Issue 3, 291-296, 2016.

Kazumi Nakamatsu K., Kountchev R. (Editors), *New Approaches in Intelligent Control. Techniques, Methodologies and Applications*, Springer International Publishing, ISBN 978-3-319-32168-4, Switzerland, 2016.

Kharrazi A., Kumar P., Saraswat C., Avtar R., Mishra B. K., "Adapting Water Resources Planning to a Changing Climate: Towards a Shift from Option Robustness to Process Robustness for Stakeholder Involvement and Social Learnings", *Journal of Climate Change*, vol. 3, no. 2, pp. 81-94, 2017.

Ramos W. B., Figueirêdo M. F., Brito K. D., Ciannella S., Vasconcelos L. G. S., Brito R. P., "Effect of Solvent Content and Heat Integration on the Controllability of Extractive Distillation Process for Anhydrous Ethanol Production", *Ind. Eng. Chem. Res.*, 55 (43), pp 11315–11328, 2016.

Roy A., Srivastava P., Sinha S., "Risk and reliability assessment in chemical process industries using Bayesian methods", *Reviews in Chemical Engineering*, 40(5), 479-499, 2014.

Tarifa E. E., Vera C. R., Franco Domínguez D., Benitez L. A., "Optimization Model for Refinery Hydrogen Networks Part I", *International Journal of Engineering Research and Applications (IJERA)*, ISSN: 2248-9622, www.ijera.com, Vol. 6.2 - Issue 10, 84-88, 2016.

Tarifa E. E., Vera C. R., Franco Domínguez D., Benitez L. A., "Optimization Model for Refinery Hydrogen Networks Part II", *International Journal of Engineering Research and Applications (IJERA)*, ISSN: 2248-9622, www.ijera.com, Vol. 6 - Part 2 - Issue 10 (October - 2016), 89-95, 2016.

Xia Z., Zhao J., "Steady-state optimization of chemical processes with guaranteed robust stability and controllability under parametric uncertainty and disturbances", *Computers & Chemical Engineering*, Volume 77, Pages 116-134, ISSN 0098-1354, 2015.

Yurquina C. D., Franco Domínguez S., Tarifa E. E., "Simulación estacionaria y optimización de un tren de pozas solares para la concentración de litio", *XI Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA*, 23 y 24 de junio, San Salvador de Jujuy, ISSN 1853-7871, 2016.

.

Extracción de Características en Audio con Redes Neuronales Convolucionales

García Mario Alejandro ✉[1], Rosset Ana Lorena[2], Eduardo Destéfanis[1]

[1]Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN FRC)
[2]Universidad 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 diseño de las capas de extracción de características de una red neuronal profunda orientada a la clasificación de la calidad vocal.

Palabras clave: ***aprendizaje automático, aprendizaje profundo , análisis acústico***

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 aprendizaje profundo (*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 aprendizaje profundo 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, se enfoca en el desarrollo de capas de extracción de características para un modelo de redes neuronales profundas.

La extracción de características es una de las etapas más importantes del aprendizaje automático. Esta tiene como objetivo la obtención de atributos del objeto a clasificar, de forma tal que un método de clasificación pueda encontrar regiones de decisión para cada una de las clases en el espacio formado por dichos atributos o características. En aprendizaje profundo, esta etapa se realiza en las primeras capas de la red neuronal.

Las medidas acústicas *shimmer*, *jitter* y HNR están relacionadas con la calidad vocal. A su vez, el cálculo de estas medidas depende del valor de la frecuencia fundamental de la voz F_0.

La estimación de F_0 es compleja y continúa siendo tema de investigación. Muchas de las técnicas de cálculo de F_0 parten de la información espectral de la señal. El espectro de frecuencias de una señal digital se obtiene con la Transformada Discreta de Fourier (DFT), por lo tanto, es deseable que la red neuronal sea capaz de calcular, en sus primeras capas, información equivalente a DFT del audio y la propague a las capas siguientes. Por ejemplo, la forma más frecuente de estimar F_0 implica el cálculo del cepstrum. Para calcular el cepstrum es necesario calcular dos veces sucesivas el *power spectrum* de la señal y esto a su vez implica dos DFT. Entonces, según el enfoque de nuestro trabajo, una red neuronal que calcule en sus primeras capas la DFT, recibirá en las capas posteriores la información necesaria para calcular F_0. A continuación se explica el cálculo de la DFT con una red neuronal.

Metodología:

Tanto la DFT como las redes neuronales, realizan transformaciones lineales. Por lo tanto, una red neuronal es capaz de calcular la DFT si sus pesos sinápticos son iguales a los coeficientes de la DFT [19].

Como para los objetivos del proyecto no es necesario preservar la información relativa a la fase en el espectro de frecuencias, se definen las capas necesarias para calcular la magnitud del espectro, es decir, el valor absoluto de la DFT. El valor absoluto de la DFT de una señal discreta $\boldsymbol{x}$ se define como:

$$|X_k| = \sqrt{\left(\sum_{n=0}^{N-1} x_n \phi(k,n)\right)^2 + \left(\sum_{n=0}^{N-1} x_n \psi(k,n)\right)^2}$$

donde N es la longitud de la secuencia $\boldsymbol{x}$, k es la frecuencia en el espectro,

$\phi(k,n) = \cos(-2\pi kn/N)$,

$\psi(k,n) = \sin(-2\pi kn/N)$.

La red neuronal que calcula la magnitud del espectro se muestra en la figura 1. Esta red devuelve un espectro de frecuencias para cada segmento de longitud N en el audio. En consecuencia a la salida de la red se obtiene un espectrograma.

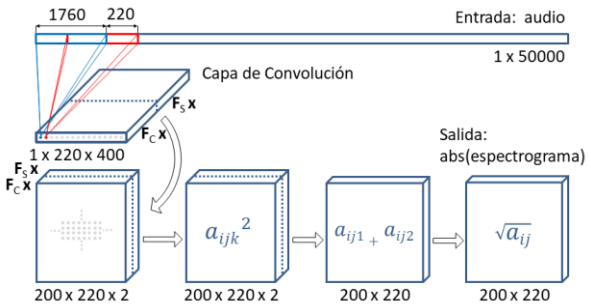


Figura 1. Modelo de neuronal de cálculo de la magnitud del espectrograma.

Los pesos de la capa de convolución son las matrices de tamaño $N \times N$, $\boldsymbol{F_C}$ y $\boldsymbol{F_S}$, que contienen los valores de $\phi(k,n)$ y $\psi(k,n)$ respectivamente. Más detalles se pueden ver en [19]. Con estos pesos, la red calcula el espectrograma de la misma forma que se hace con la DFT. La ventaja de hacerlo con una red neuronal, es que, si los pesos se pueden aprender, la red puede adaptarse a problemas particulares para los que exista un set de pesos mejores que los de $\phi(k,n)$ y $\psi(k,n)$. Para comprobar si la red tiene capacidad de adaptación, se realiza el entrenamiento partiendo de pesos aleatorios. Las entradas utilizadas son audios reales de vocales sostenidas de dos segundos de duración y las

salidas esperadas son espectrogramas calculados con la DFT.

3. RESULTADOS OBTENIDOS

A continuación se exponen los resultados de 30000 ciclos de entrenamiento del modelo propuesto. Los pesos se inicializaron con valores aleatorios entre -10^{-6} y 10^{-6} uniformemente distribuidos. Se utlizó el método de optimización Adam [20] con los parámetros provistos por los autores y las actualizaciones de los pesos se realizaron en batchs de tamaño 300 (la totalidad de datos de entrenamiento).

Los cálculos se realizaron sobre una GPU NVIDIA Titan Xp, donada a través del GPU Grant Program de NVIDIA.

El MSE de validación alcanzado fue 1.41×10^{-6} (9.79×10^{-6}% del valor medio de la salida esperada), mientras que para el mismo modelo con los pesos asignados de forma directa se logra un MSE < 10^{-9} [19].

Estos resultados indican que el modelo se puede entrenar y, por lo tanto, también puede adaptarse a problemas particulares.

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 Phoniatrica 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., Puhrsch, 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] García, M.A., Destéfanis, E.A.: Spectrogram Prediction with Neural Networks. XXIV Congreso Argentino de Ciencias de la Computación (2018).

[20] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)

Framework SDF Machine Learning en transacciones financieras y detección temprana de fraudes

Fabian Frola[1], Carlos Iván Chesñevar[2], Carlos Alvez[1], Graciela Etchart[1], Ernesto Miranda[1],Silvia Ruiz[1], Juan José Aguirre[1], Juan Carlos Teze[1,2]

[1] Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos

Av. Tavella 1424 (E3202KAC), Concordia, Entre Ríos, Argentina.

[2] Instituto de Ciencias e Ingeniería de la Computación (ICIC CONICET UNS)

Universidad Nacional del Sur, San Andrés 800, 8000 Bahía Blanca, Argentina.

fabianenlared@gmail.com, cic@cs.uns.edu.ar, {caralv,getchart, emiranda, sruiz, juaagu}@fcad.uner.edu.ar, jcarlt02@gmail.com

Resumen

En la actualidad, con el crecimiento exponencial de transacciones financieras de tarjetas de crédito y débito, la disminución de barreras de acceso, la globalización y la inclusión financiera se ha incrementado en mayor medida el fraude y la inteligencia creativa para la mutación del comportamiento fraudulento. Es de vital importancia la detección temprana de fraude aplicando distintas estrategias basadas en inteligencia artificial que puedan mitigar, disminuir, y prevenir este flagelo. El objetivo de este trabajo es estudiar, analizar los fundamentos, técnicas, estrategias y herramientas de machine learning que nos permitan dar el paso necesario para abordar el tema de autorizaciones financieras y detección de fraude, cuyo abordaje se hace inalcanzable con estrategias determinísticas o algoritmia tradicional. A partir del estudio mencionando se construirá un *framework* consolidado aplicable en cada etapa del proceso, desde la adquisición de datos, tanto en línea como históricos, el pre-procesamiento, la clasificación y los aportes al modelo predictivo para la detección de fraude.

Palabras clave: machine learning, inteligencia artificial, detección de fraude financiero.

Contexto

Esta línea de trabajo está siendo desarrollada en conjunto por la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos y el Instituto de Ciencias e Ingeniería de la Computación de la Universidad Nacional del Sur (ICIC CONICET UNS); los resultados vinculados al mismo se enmarcan en una tesis de la Maestría en Sistemas de Información (en desarrollo), dictada en la FCAD-UNER, y en el trabajo interdisciplinar con el grupo de investigación relacionado al proyecto PID 07/G044 (donde el foco radica en la identificación de personas a partir de

rasgos biométricos), dando continuidad a la aplicabilidad de modelos predictivos en nuevos contextos.

Introducción

En nuestra experiencia en el mundo financiero nos hemos encontrado con un crecimiento exponencial de transacciones, así como de los fraudes asociados a ellas, lo que incrementa la complejidad a la hora de detectarlos tempranamente como sería deseable. Por otro lado, se aprecia que la banca y todo el conjunto de entidades financieras que forman parte de este escenario se ha regido tradicionalmente por modelos algorítmicos tradicionales (determinísticos). Esta situación, si bien parece deseable, es en realidad el principal obstáculo tanto a la hora de autorizar una transacción de crédito/débito, así como la capacidad para detectar un eventual fraude, donde suele haber información incompleta y potencialmente inconsistente, lo que obliga a la construcción de modelos utilizando técnicas de inteligencia artificial a través de Machine Learning.

La toma de decisiones utilizando reglas de negocio pautadas de antemano (que muchas veces carecen de justificaciones racionales) representa un factor negativo tanto en la satisfacción del cliente así como la capacidad de aprendizaje del propio modelo de lo que está ocurriendo en la realidad, limitando las adecuaciones necesarias para acompañar la globalización y la diversificación de la economía, factores con un alto impacto a nivel bancario.

La economía digital se ha consolidado con un crecimiento exponencial en los últimos años, sumado a la expansión de los modelos de pagos electrónicos, transacciones, tarjetas de crédito y débito, e-commerce. El fraude ha acompañado este crecimiento exponencial, generando costos de varios miles de millones de dólares anuales, tanto en pérdidas de los individuos como para las organizaciones. Si bien desde el punto de vista del fraude financiero su aparición suele ser menor al 1/100 de las transacciones (y muchas veces el costo en sí de una acción fraudulenta no justifica el costo para detectarla oportunamente), el costo intrínseco del fraude no se puede valorar solo desde una óptica monetaria, sino que también intervienen otras variables como ser la fidelización de clientes y la erosión de la confianza (asociada a la permanencia o pérdida de clientes). El diseño de algoritmos de detección de fraudes es particularmente desafiante en el ámbito financiero de transacciones con tarjetas de crédito y débito, debido a la distribución no estacionaria de los datos, la clasificación altamente desequilibrada y disponibilidad de pocas transacciones etiquetadas por los investigadores de fraude.

El fraude en el sector financiero afecta la economía de individuos, organizaciones y naciones. Se estima que más de una de cada tres organizaciones han sufrido cualquier tipo de fraude en los últimos dos años. Pero es aún más sorprendente que alrededor del 10% de los delitos financieros se detectan esencialmente por casualidad [1]. En el ámbito financiero, históricamente muy conservador, muy pocos se arriesgan a adoptar criterios poco determinísticos a la hora de decidir si autorizar o negar una transacción financiera; lo mismo ocurre en la detección asociada a un fraude, que generalmente pasa a ser de índole forense, *a posteriori* de la ocurrencia del mismo y sin un registro sistemático (lo que permitiría realizar trabajo de inteligencia sobre quienes cometen el fraude). Más allá de esto, es común que una vez detectado el modus operandi por el sistema bancario generalmente el criminal tiende a cambiar la forma de realizarlo. El desequilibrio relativo del fraude en un conjunto de datos puede comenzar en con la proporción de un defraudador cada 1000 clientes normales, por lo que se trasforma en un problema cuya detección es extremadamente compleja. Si

tomamos los algoritmos de clasificación de minería de datos para tratar de clasificarlos y lograr un modelo predictivo aplicable a futuras ocurrencias se ha llegado a verificar tasas de acierto de hasta 99% [4], pero esto no es suficiente dado que la identificación de falsos positivos deriva a menudo en mayores problemas (como ser la pérdida de un cliente asiduo).

En este trabajo confluyen varias líneas de investigación, entre las que se destacan la aplicación de minería de datos [5] [6] y las técnicas de clasificación [1] [8] [9] [10] y predicción (con las particularidades que presenta un domino de aplicación en el cual aparece una clase ampliamente predominante -clientes normales- y en contraposición muy pocos eventos a clasificar en forma negativa (esto es, los fraudes) [2].

También debemos reconocer que cualquier sistema predictivo de fraude deberá tener un conjunto de datos de entrenamiento apropiado para generar el modelo, y esto representa una de las grandes barreras a la investigación dado que generalmente los datos a utilizar son privativos a algunas pocas organizaciones que tienen acceso (estando protegidos por leyes que amparan la privacidad de los mismos). Sin perjuicio de esto último resulta relevante abordar el desempeño integral de un clasificador "minoritario" (como se sugiere en [3]) que permita consolidar a los modelos predictivos como la alternativa para detección de fraude o previsión de situaciones que conllevan acciones fraudulentas.

Cabe destacar que la amplia adopción del estándar ISO-8583 [7] en el sistema financiero uniformiza la adquisición de datos de transacciones que permitan la captura de información relevante (necesaria como datos de entrada para la generación de modelos predictivos basados en el aprendizaje automatizado).

Líneas de Investigación y Desarrollo

Una de las principales líneas de investigación de este trabajo es la aplicación de machine learning en la detección temprana de fraude de transacciones financieras con tarjetas de crédito / débito, realizando una revisión actualizada de las investigaciones en el estado del arte. Se busca así avanzar hacia la construcción de un *framework* SDF (Sistema Detección de Fraude).

Uno de los principales objetivos de SDF será el reconocimiento de transacciones fraudulentas con elevados niveles de certeza, en un contexto especialmente complejo por el gran desbalanceamiento de las clases (contrastando la clase minoritaria de fraudes frente al universo mayoritario de transacciones financieras genuinas).

Cabe señalar que cada individuo tiene características diferentes de comportamiento en un contexto de no estacionalidad de los datos. Por ende no resulta sencillo realizar procesos de detección de fraude en este ámbito: hay una alta sensibilidad a los impactos negativos que ocasionen los falsos positivos (que erosionan la confianza entre el cliente y la institución financiera), lo que hace que muchas veces es deseable para las mismas padecer los costos directos del fraude y no afectar al cliente. Al mismo tiempo, también es apreciable que la masificación de la inclusión financiera, así como los impactos de la globalización que han profundizado las consecuencias de la inadecuda gestión del fraude.

Resultados y Objetivos

El objetivo es identificar las mejoras aplicables a la clasificación de datos en la clase minoritaria, en entornos de transacciones financieras, para la minimización de falsos positivos. Con este fin se está desarrollando una arquitectura (Fig. 1) con capacidades predictivas de

situaciones de fraude, seleccionando técnicas de machine learning aplicables a las autorizaciones financieras y detección de fraudes en la generación de modelos predictivos, combinando el enfoque supervisado y no supervisado.

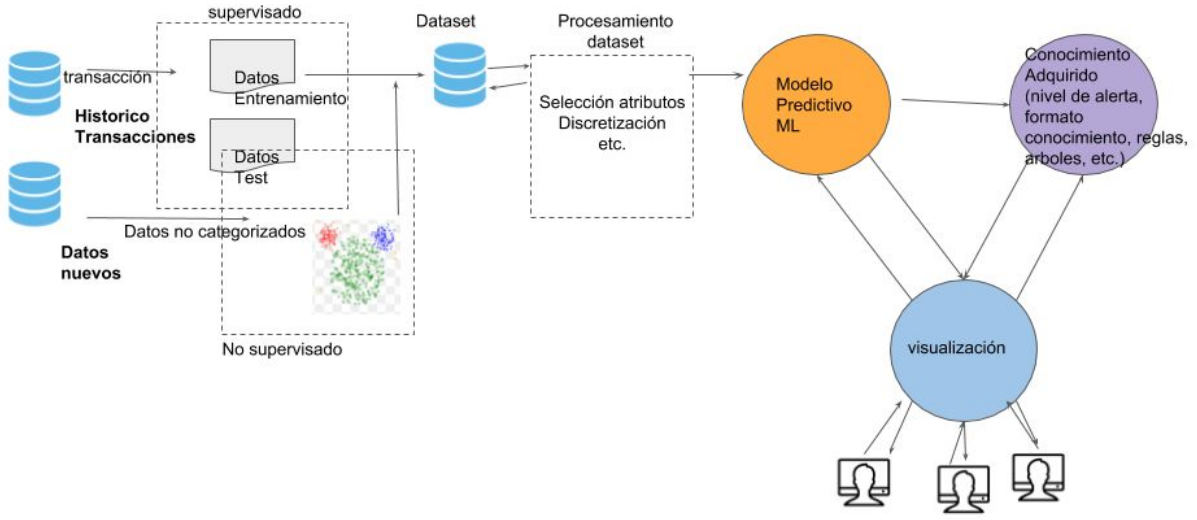


Fig. 1: Arquitectura *framework* Propuesta

Formación de Recursos Humanos

La presente línea de investigación se lleva adelante por un equipo de integrantes de la FCAD/UNER y el ICIC CONICET UNS (Bahía Blanca, Argentina). Se desprende que esta sinergia favorece la formación de recursos humanos en el ámbito de la inteligencia artificial. La línea de investigación involucra varios docentes investigadores y cinco tesistas de maestría. En este contexto se está realizando una tesis de Maestría (orientada a la caracterización de la arquitectura SDF), brindándose también un aporte significativo en otras líneas de trabajo del grupo de investigación..

Referencias

[1] Stefan Axelsson Edgar Alonso Lopez-Rojas. A review of computer simulation for fraud detection research in nancial datasets. IEEE, 2017

[2] Fabrizio Carcillo, Yann-Ael Le Borgne, Olivier Caelen, y Gianluca Bontempi. Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization. I. J. Data Science and Analytics, 5(4):285-300, 2018. doi:10.1007/s41060-018-0116-z. URL https://doi.org/10.1007/s41060-018-0116-z.

[3] Michele Carminati, Alessandro Baggio, Federico Maggi, Umberto Spagnolini, y Stefano Zanero. Fraudbuster: Temporal analysis and detection of advanced financial frauds. En Cristiano Giuffrida, Sébastien Bardin, y Gregory Blanc, eds., Detection of Intrusions and Malware, and Vulnerability Assessment, págs. 211-233. Springer International Publishing, Cham, 2018. ISBN 978-3-319-93411-2.

[4] Andrea Dal Pozzolo. Adaptive Machine Learning for Credit Card Fraud Detec-
tion. Tesis Doctoral, Machine Learning Group, Computer Science Department. Universite Libre de Bruxelles, 2015.

[5]Alex Guimaraes Cardoso de Sá, Adriano C. M. Pereira, y Gisele L. Pappa. A customized classification algorithm for credit card fraud detection. Eng. Appl. of AI, 72:21-29, 2018. doi:10.1016/j.engappai.2018.03.011. URL https://doi.org/10.1016/j.engappai.2018.03.011.

[6] Wei Dong, Shaoyi Liao, y Zhongju Zhang. Leveraging financial social media data for corporate fraud detection. Journal of Management Information Systems, 35(2):461-487, 2018. doi:10.1080/07421222.2018.1451954. URL https://doi.org/10.1080/07421222.2018.1451954.

[7] ISO8583-1 Financial transaction card originated messages -- Interchange message specification. Part 1: Messages, data elements and code value. ISO, 2003. URL https://www.iso.org/obp/ui/#iso:std:iso:8583:-1:ed-1:v1:en

[8] Dongxu Huang, Dejun Mu, Libin Yang, y Xiaoyan Cai. Codetect: Financial fraud detection with anomaly feature detection. IEEE Access, 6:19161-19174, 2018. doi:10.1109/ACCESS.2018.2816564. URL https://doi.org/10.1109/ACCESS.2018.2816564.

[9] Rafiq Ahmed Mohammed, Kok Wai Wong, Mohd Fairuz Shiratuddin, y Xuequn Wang. Scalable machine learning techniques for highly imbalanced credit card fraud detection: A comparative study. En PRICAI 2018: Trends in Artificial Intelligence - 15th Pacific Rim International Conference on Artificial Intelligence, Nanjing, China, August 28-31, 2018, Proceedings, Part II, págs. 237-246. 2018. doi:10.1007/978-3-319-97310-4n 27. URL https://doi.org/10.1007/978-3-319-97310-4_27.

[10] Kuldeep Randhawa, Chu Kiong Loo, Manjeevan Seera, Chee Peng Lim, y Asoke K. Nandi. Credit card fraud detection using adaboost and majority voting. IEEE Access, 6:14277-14284, 2018. doi:10.1109/ACCESS.2018.2806420. URL https://doi.org/10.1109/ACCESS.2018.2806420.

Hacia ciudades mas eficientes y sostenibles

Villagra A.[1], Errecalde M.[1,2], Pandolfi D[1]., Molina D. [1], Varas V[1]., Orozco S. [1], Valdéz J. [1], Rasjido J. [1], Mercado V[1]., Carballo L. [1], Pérez D.[1], Montenegro C.[1]

[1]Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada (ITA) - 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

avillagra@uaco.unpa.edu.ar, merreca@unsl.edu.ar, {dpandolfi, dmolina, vvaras, sorozco, jcvaldez, jrasjido, vmercado,lcarballo, dperez,cmontenegro}@uaco.unpa.edu.ar

RESUMEN

Los nuevos retos de la ciudad y las limitaciones, tanto de nuestro entorno, como de los recursos de todo tipo que necesita de forma continuada, en la mayoría de los casos provenientes de fuentes finitas (energéticos, infraestructuras, económicos, etc.), hace que la disposición de éstos, su utilización y empleo de la forma más eficiente posible sea una de las claves si se quiere conseguir niveles adecuados de sostenibilidad. La Ciudad Inteligente (*Smart City* o SC) ha de ser capaz de dar respuesta a la necesidad de las ciudades de ser más eficientes y sostenibles, en aras de la consecución de un verdadero desarrollo urbano sostenible.

Se puede decir que una SC es la visión holística de una ciudad que aplica las Tecnologías de la Información y Comunicación (TIC) para la mejora de la calidad de vida y la accesibilidad de sus habitantes y asegura un desarrollo sostenible económico, social y ambiental en mejora permanente.

Esta línea de trabajo se presenta como un aporte en el camino hacia el desarrollo sostenible de una cuidad inteligente. En particular, en este proyecto se propone identificar, resolver y desarrollar prototipos de aplicaciones y servicios en tres ejes movilidad, sociedad y entorno.

Palabras clave: Ciudades Inteligentes, Metaheurísticas, Big Data, Transito, Residuos tecnológicos.

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/B225 "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).

1. INTRODUCCIÓN

Según la Conferencia de las Naciones Unidas sobre el Desarrollo Sostenible, celebrada en Río de Janeiro en 2012, la mitad de la humanidad vive en ciudades. La población urbana ha aumentado desde los 750 millones de personas –en 1950–, hasta los 3.600 millones en 2011. Se estima que hacia 2030 casi un 60% de la población mundial residirá en zonas urbanas. Este notable crecimiento supondrá que, en los próximos

años, ciudades de todo el mundo deberán prestar servicios, fundamentalmente no administrativos, de manera sostenible. Atender a las necesidades de agua potable, de transporte, o de aire limpio, constituirá un reto científico de extraordinaria magnitud y una oportunidad para la industria [CDN09]. 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.

Según [GF+07] una ciudad inteligente es aquella que realiza actividades en al menos una de las seis áreas de acción inteligente: *Smart Economy*, *Smart People*, *Smart Governance*, *Smart Mobility*, *Smart Environment* y *Smart Living*. La definición coincide con las ideas de B. Cohen [Coh12], quien en 2012 propuso una rueda donde se integran los ejes mencionados con otras actividades para su puesta en marcha. A pesar de estos avances en la definición, conocimiento y aparente penetración del concepto de ciudades inteligentes, este dominio apenas está despegando.

Desde el punto de vista de datos, la expansión de Big Data y la evolución de las tecnologías de IoT (*Internet of Things*) 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].

Dada la proliferación de varios dispositivos como RFIDs, sensores y actuadores IoT está creciendo también. Los dispositivos inteligentes (dispositivos que tienen capacidades computacionales significativas, transformándolos a "Cosas inteligentes") están integradas en el entorno para monitorear y recopilar información ambiental en una ciudad inteligente.

A medida que el paradigma de IoT se va extendiendo, las 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.

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]).

Crear aplicaciones inteligentes en una ciudad inteligente es uno de los mayores desafíos a los que podemos enfrentarnos. Los problemas en las ciudades inteligentes son variados (economía, movilidad, gobernanza, personas, vida y entorno), multidisciplinares por naturaleza, aparentemente inconexos y requieren un vasto aparato científico. 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.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta sección se describen las líneas de investigación que se llevan 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 in-

teligente, 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: movilidad, sociedad y entorno.

Con respecto a la movilidad inteligente, se ha comenzado a trabajar problemas de tráfico y en particular en el control de ciclos semafóricos. Los problemas de tráfico aumentan constantemente y las ciudades del futuro solo pueden ser verdaderamente inteligentes si habilitan *Smart Mobility*. Aplicaciones de movilidad inteligente como el control de semáforos, el estacionamiento inteligente y la gestión del tránsito, entre otras. están comenzando a ser utilizadas en todo el mundo, trayendo beneficios a las ciudades, mejor calidad de vida, costos reducidos, uso de energía más eficiente y disminución de las emisiones vehiculares. Mejorar la movilidad representa un gran reto.

Actualmente, una ciudad con tráfico congestionado y muchos embotellamientos conduce a más contaminación. Por lo tanto, con la gestión inteligente del tráfico, el consumo de combustible y las emisiones contaminantes pueden ser reducidas. A diferencia de técnicas tales como semáforos inteligentes que requieren nueva infraestructura, ubicación de sensores y modificaciones en obras civiles ya construidas, el uso de técnicas de inteligencia artificial en la optimización de los ciclos de semáforos se presenta como una herramienta viable, rápida, eficiente y de bajo costo [GF+14]. En esta línea nos centramos en dos versiones de un Algoritmo Genético Celular (cGA)[AB09], aplicado al problema de planificación de los ciclos de los semáforos. Las soluciones obtenidas por los cGAs son simuladas por el popular micro-simulador SUMO (*Simulator of Urban Mobility*) [KB+06]. Se evalúan grandes escenarios cercanos a la realidad de vastas áreas urbanas.

En cuanto a sociedad inteligente, se está trabajando en temas de *Big Data* y Minería de Datos en análisis de autoría (AA) [Sta09], en particular la determinación del perfil del autor (DPA), es decir, aquella que identifica patrones compartidos por un grupo de gente y que aborda problemas de clasificación de los usuarios de la Web de acuerdo a la edad, género, orientación política, etc. La DPA, un sub-campo del área más general conocida como análisis de autoría (AA), es un tema muy importante de investigación por sus potenciales (y actuales) aplicaciones en problemas de seguridad nacional e inteligencia, lingüística forense, análisis de mercados e identificación de rasgos de personalidad, entre otros. Particularmente, se ha tomado como caso de estudio primario los documentos de periodistas con diversas orientaciones políticas (oficialista vs opositor) con el objetivo de realizar con los mismos el Análisis de Autor y la Determinación/Caracterización del perfil del autor.

Por otro lado, en cuanto al entorno inteligente se está trabajando en la gestión adecuada de los residuos urbanos y en particular los residuos tecnológicos. Este tema es de creciente preocupación para los agentes públicos, con impactos directos sobre la salud, el medio ambiente y la calidad de vida de la población [AZ+17]. 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. La recolección inteligente de desechos electrónicos no solo es relevante en términos de impacto ambiental; también hay un componente económico muy importante. El valor de las materias primas presentes

en los desechos electrónicos y el aprovechamiento de estos elementos creando una economía circular al revalorizarlas.

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 ejes de desarrollo en ciudades inteligentes permitirá hacer prototipos reales, aprender más y producir un mayor impacto en Ciencia e Industria. Dada la amplia visión se espera generar oportunidades para nuevos proyectos y socios.

En particular, se han obtenido resultados promisorios con nuestros algoritmos aplicados al problema de *scheduling* de ciclos de luces de tráfico. La performance de nuestros enfoques mejora a técnicas paralelas del estado del arte. Como trabajo futuro se intenta reducir el costo computacional en términos de tiempo de procesamiento aplicando funciones subrogadas. de las evaluaciones de las soluciones a través de la incorporación de funciones. Además, aplicando el conocimiento obtenido de los trabajos realizados se pretende ajustar la función de evaluación de las soluciones a fin de explorar más eficientemente el espacio de búsqueda y obtener una mayor reducción en cuanto a la emisión de gases.
Además, se espera realizar la publicación de los resultados del proyecto en revistas de alto impacto y conferencias de 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

El equipo de trabajo se encuentra formado por dos Doctores en Ciencias de la Computación, dos Magisters en Ciencias de la Computación, cinco Ingenieros en Sistemas y tres estudiantes de la Carrera Ingeniería en Sistemas de la UNPA.

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 dos becarios de grado y un becario de posgrado.

5. BIBLIOGRAFIA

[AB09] Alba, E., & Dorronsoro, B. (2009). Cellular genetic algorithms (Vol. 42). Springer Science & Business Media.

[AZ+17] Anagnostopoulos, T., Zaslavsky, A., Kolomvatsos, K., Medvedev, A., Amirian, P., Morley, J., & Hadjieftymiades, S. (2017). Challenges and opportunities of waste management in IoT-enabled smart cities: a survey. IEEE Transactions on Sustainable Computing, 2(3), 275-289.

[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.

[CD+09] Caragliu A., Del Bo C., Nijkamp P., “Smart Cities in Europe”, Proceedings of the 3rd Central European Conference in Regional Science – CERS, pp.45-59, 2009

[Coh12] Cohen Boyd, What Exactly Is A Smart City? http://www.fastcoexist.com/1680538/what-exactly-is-a-smart-city, 2012

[GF+07], Giffinguer, R.; Fertner, C; Kramar, H.; Meijers, E. (2007) Ranking of European Medium-Size Cities. Centre of Regional Science, Univ. Tecnológica de Viena.

[GF+14] Garcia-Nieto, J., Ferrer J., & Alba E. (2014). Optimising traffic lights with metaheuristics: Reduction of car emissions and consumption. In Neural Networks (IJCNN), 2014 International Joint Conference on, pages 48-54. IEEE, 2014.

[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., Bousselham, 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.

[KB+06] Krajzewicz, D., Bonert, M., & Wagner, P. (2006). The open source traffic simulation package SUMO. RoboCup 2006 Infrastructure Simulation Competition, 1, 1-5.

[Kun14] Kunzmann, K. R. (2014). Smart cities: a new paradigm of urban development. Crios, 4(1), 9-20.

[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.

[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.

[Sta09] Stamatatos, E. (2009). A survey of modern authorship attribution methods. Journal of the American Society for information Science and Technology, 60(3), 538-556.

[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.

Herramientas para el desarrollo de sistemas de análisis de textos no estructurados

Marina Cardenas[1], Julio Castillo[1], Martin Navarro[1], Nicolás Hernández[1], Melisa Velazco[1]
[1] Laboratorio de Investigación de Software/Dpto. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{ing.marinacardenas, jotacastillo}@gmail.com

Resumen

En este artículo se describen las actividades desarrolladas y los subsistemas que conforman el proyecto de investigación que se denomina Desarrollo de Sistemas de Análisis de Texto, un proyecto de investigación homologado por la Secretaría de Ciencia y Técnica (SCyT) de la UTN.

En este proyecto se trabaja en el desarrollo de herramientas que permitan realizar el análisis de información textual de una manera más eficiente, e involucra la creación de material de entrenamiento de los sistemas de análisis de textos y el desarrollo de herramientas software que sirvan para el análisis y procesamientos de grandes volúmenes de textos.

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

Contexto

Este artículo presenta el proyecto denominado Desarrollo de Sistemas de Análisis de Texto (ADT), que es un proyecto homologado por la SCyT de la UTN, que se enmarca dentro del área de lingüística computacional, y que tiene como objetivo el desarrollo de herramientas para análisis textual de diversas fuentes de información en formato no estructurado.

Actualmente, el proyecto se encuentra dentro del grupo de investigación denominado Grupo de Aprendizaje Automático, Lenguajes y Autómatas (GA^{2}LA) con fecha de creación del grupo en Octubre de 2018 en la UTN-FRC.

En el grupo GA^{2}LA se desarrollan proyectos relacionados con autómatas y lenguajes formales, procesamiento del lenguaje natural, y aprendizaje automático, y en especial proyectos orientados a la aplicación de la inteligencia artificial para resolver problemas de las ciencias sociales.

En este contexto, los recursos humanos con los que cuenta el grupo son ingenieros en sistemas de información, licenciados/doctores en ciencia de la computación, demógrafos y arquitectos, y se complementa con becarios y pasantes.

Los problemas que aborda este proyecto de investigación están relacionados a la extracción de información, detección y reconocimiento de paráfrasis, y de reconocimiento de implicación textual. Es decir, problemas relacionados a la identificación de oraciones (o párrafos) que tengan el mismo significado, o bien la identificación de oraciones-párrafos que estén semánticamente relacionados entre sí mediante una relación de implicación.

Físicamente, los integrantes del proyecto desarrollan sus actividades en el Laboratorio de Investigación de Software LIS[1] del Dpto. de Ingeniería en Sistemas de Información de la Universidad Tecnológica Nacional Facultad Regional Córdoba (UTN-FRC).

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

1. Introducción

Los sistemas de análisis de texto se enfrentan a problemas difíciles dentro del área de la ciencia de la computación, debido principalmente a la dificultad del análisis del lenguaje (derivada de la ambigüedad del lenguaje) relacionado a la etapa de análisis semántico, como así también, a los relativamente escasos materiales de entrenamiento y a la capacidad de cómputo necesaria para correr determinados algoritmos muy demandantes en recursos de hardware.

Como una manera de abordar el problema de la escasez del material de entrenamiento, en este proyecto se trabaja con la creación de corpus lingüísticos que puedan ser utilizados como material de entrenamiento en sistemas de aprendizaje supervisado o semi-supervisado. Una vez construidos, los conjuntos de entrenamiento pueden ser de utilidad en diversos problemas relacionados con el análisis de textos.

Adicionalmente, la correcta construcción y diseño del material de entrenamiento puede ayudar a abordar el problema de la ambigüedad en el texto, que son complemento de diversas técnicas de análisis del discurso.

En cuanto a la capacidad de cómputo, es un desafío pendiente, y trabajo futuro, el paralelizar algoritmos secuenciales que necesitan gran poder de cómputo y que son costos en cuanto al tiempo de ejecución.

Este 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, entre ellas destacamos las basadas en redes neuronales artificiales [6][7][8], máquinas kernel [9], deep learning [10][11], y árboles de decisión.

En el marco de este proyecto se están desarrollando sistemas de análisis y procesamiento de texto, entre los que podemos destacar:

- 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.
  Este software ha permitido la creación de corpus para sistemas de reconocimientos de paráfrasis.

- 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.
  Este software permite centralizar en un repositorio común, la información dispersa en grandes bases de datos.

- Sistema de detección de similitudes en archivos de código fuente (SDS). Es un sistema que tiene como objetivo determinar la similitud de archivos de código fuente escritos en diferentes lenguajes de programación. Se está trabajando a nivel de granularidad de archivos de código fuente y a nivel de funciones o módulos.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de investigación de este proyecto es la lingüística computacional abordada con técnicas de aprendizaje automático.

Como se ha mencionado anteriormente, una de las herramientas desarrolladas es un software Asistente de Creación de Corpus, el cual forma parte de la línea de investigación denominada Lingüística de Corpus [12]. Esta rama de la lingüística se caracteriza porque los resultados de las investigaciones se realizan en base a las

entidades lingüísticas obtenidas de textos tabulados y clasificados.

El equipo de investigación también trabaja en el desarrollo de modelos computaciones en general. Uno de los modelos desarrollados permite modelar la ocurrencia de incendios forestales, y forma parte de otro proyecto de investigación que se denomina Modelos para la Predicción de Incendios Forestales, un proyecto homologado por la SCyT de la UTN.

La línea de innovación está relacionado a la integración de los resultados obtenidos de la lingüística de corpus, de la construcción de modelos y del uso de técnicas de aprendizaje automático, en el marco del grupo GA^{2}LA [13].

3. Resultados

Los subsistemas que se mencionan en la Sección de Introducción se están desarrollando en paralelo, y son los siguientes:

- Asistente de Creación de Corpus (ACC),
- Sistema de Mapeo de Datos (SMD), y
- Sistema de Detección de Similitudes en archivos de código fuente (SDS).

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.

El ACC es una herramienta semiautomática que permite a los usuarios sistematizar e identificar los fenómenos lingüísticos presentes en diversos textos. Además, permite clasificar pares de texto con paráfrasis, a la vez que facilita la lectura y el estudio de los corpus generados.

Como resultado se generan corpus etiquetados que son necesarios en la etapa de entrenamiento en sistemas de aprendizaje supervisado.

Esta herramienta permite registrar diversos fenómenos lingüísticos a nivel léxico, sintáctico, morfológico y semántico. El material de entrenamiento construido es para el idioma español e inglés, y es utilizado en sistemas de RTE (Implicación Textual).

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.

Se han creado dos corpus monolingües en español con 100 pares de elementos [14], clasificados según se describe en [15]. Además, se está terminado de construir un tercer corpus de 100 pares monolingüe en inglés.

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.

El SMD está basado en una aplicación web que almacena los datos normalizados en una estructura estándar de una base de

datos SQL Server facilitando la búsqueda y análisis de textos.

La aplicación permite la selección del origen y destino de datos estructurados, y para realizar el mapeo (transformación) de datos de manera interactiva.

Una vez construido un repositorio para un dominio de problema dado, es posible realizar minería de datos sobre el mismo.

El tercer subsistema se denomina Sistema de Detección de Similitudes en códigos fuente (SDS). Es el subsistema de desarrollo más reciente, y hasta el momento, permite la integración de diversas medidas de similitud léxica aplicadas sobre archivos de códigos fuentes escritos en el mismo lenguaje de programación (C o Java) [16].

La identificación de similitudes de código puede servir para varios propósitos, como la trazabilidad en proyectos de desarrollo de software, detección de reutilización de código, identificación de vulnerabilidades en el código, y detección de plagio.

El sistema SDS permite efectuar una comparación de un archivo de código fuente contra un conjunto de archivos. Debe notarse que la identificación de dos archivos potencialmente similares requiere un costo cuadrático en cantidad de comparaciones. Estas operaciones son muy costosas cuando el tamaño del conjunto es grande. Es por esta razón que se necesita técnicas de multiprocesamiento para obtener resultados en tiempos aceptables. Esto es parte de nuestro trabajo futuro.

4. Formación de Recursos Humanos

El equipo de investigación está formado por docentes, alumnos y egresos de la Universidad Tecnológica Nacional, Facultad Regional Córdoba, que a continuación se detallan:

- Una doctorando en ingeniería con mención en sistemas de información en la UTN-FRC, que está trabajando específicamente en el subsistema de detección de similitudes en archivos de código fuente (SDS). Además, realiza la dirección de becarios de posgrado y de becarios de grado en el contexto del proyecto.
- Un doctor en ciencias de la computación, cuya tarea principal es la dirección del proyecto, y la dirección de becarios y prácticas supervisadas.
- Un maestrando en Ingeniería en Sistemas de Información de la UTN-FRC.
- Anualmente participan en el proyecto, alumnos que realizan su práctica supervisada. La cantidad de alumnos depende de la necesidades del proyecto y de los alumnos dispuestos a trabajar en el proyecto. Las prácticas supervisadas son un requisito necesario para la obtención del grado de Ingeniero. Las mismas pueden realizarse en la industria o en el ámbito académico en el contexto de proyectos de investigación homologados.
- En el proyecto participan dos alumnos becarios anualmente, complementando así su formación curricular desde el punto de vista científico.
- Adicionalmente, se realizan 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 Icetal 2010, LNCS, vol. 6233, pp.97-102, 2010.

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

[7] Lewis, D.. Evaluating and optmizing 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. *F*undamentals of Deep Learning*:* Designing Next-Generation Artificial Intelligence Algorithms. O' Really book. 2015.

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

[13] Vázquez, Juan C., Castillo, Julio J., Constable, Leticia, Cardenas, Marina E. GA²LA: Grupo de Aprendizaje Automático, Lenguajes y Autómatas. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). 2018.

[14] Cardenas Marina E., Castillo Julio J., Navarro Martín, Hernández Nicolás A., Velazco Melisa. Sistemas de análisis textual en formato no estructurado. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). 2018.

[15] 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. CONAIISI 2017.

[16] Cardenas, Marina E., Castillo, Julio J. Procesamiento de textos estructurados. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). 2018.

Implementación de herramientas computacionales para el manejo de preferencias en servicios de razonamiento rebatible

Graciela Etchart[1], Juan Carlos Teze[1,2], Carlos Alvez[1], Guillermo R. Simari[2]

[1]Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos
Av. Tavella 1424, Concordia, Entre Ríos - CP 3200

[2]Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur
San Andrés 800 - Altos de Palihue, (8000) Bahía Blanca, Argentina

e-mail: [1]{getchart,caralv}@fcad.uner.edu.ar, [2]{jct,grs}@cs.uns.edu.ar

Resumen

Como un mecanismo práctico para formalizar el razonamiento de sentido común, los sistemas argumentativos han mostrado su potencial para aplicaciones en diversas áreas. En este sentido, el rol del criterio de preferencia en estos sistemas, el cual es utilizado para comparar argumentos en conflicto, es central para el proceso de respuesta a las consultas del usuario. Si el criterio utilizado no es adecuado al dominio representado, esto puede llevar a que las respuestas del sistema fallen al caer en la mayoría de los casos en un estado de indecisión. En este contexto, el objetivo general de este trabajo es avanzar en la implementación de diferentes herramientas computacionales para el manejo de múltiples criterios de preferencia, de manera tal que sea posible manipularlos de forma dinámica. Como resultado se buscará comprobar empíricamente la correctitud, computabilidad y eficiencia de dichas implementaciones.

Palabras clave: Manejo de Preferencias, Argumentación Rebatible, Sistemas Recomendadores

Contexto

Este trabajo se da en el marco del Proyecto PID *"Modelos de Machine Learning para la mejora de la precisión, seguridad y eficiencia en la gestión de datos biométricos"*, que da continuidad a los Proyectos PID 07/G035 "*Identificación de personas mediante Sistemas Biométricos. Estudio de factibilidad y su implementación en organismos estatales*" y PID 07/G044 "*Gestión de datos biométricos en base de datos objeto - relacionales*" [1, 2, 3].

Además, este trabajo se realiza en el marco del desarrollo de una tesis para la Maestría en Sistemas de la Información de la Universidad Nacional de Entre Ríos.

Introducción

Aquellos contextos donde la información que se maneja es incompleta o potencialmente contradictoria constituyen un escenario ideal para los sistemas argumentativos [4]. El mecanismo de inferencia sobre el cual están basados, permite decidir entre conclusiones contradictorias y adaptarse fácilmente a entornos cambiantes. Para obtener una respuesta, el proceso de razonamiento argumentativo involucra una serie de

etapas [5]. Entre ellas, es de gran importancia la comparación de argumentos en conflicto para decidir qué argumento prevalece; esto requiere una comparación o un criterio de preferencia definido para el conjunto de argumentos.

En los sistemas argumentativos actuales la definición del criterio de comparación de argumentos sigue uno de los siguientes dos enfoques: el criterio es fijo y está embebido en el sistema, o es modular. Independientemente del enfoque considerado, si el criterio no es adecuado al dominio representado, incrementan las posibilidades de indecisión cuando dos argumentos son comparados, haciéndolo menos efectivo; afectando en consecuencia la confianza del usuario en las respuestas que éste pueda dar. En este sentido, un importante desafío para avanzar es considerar el desarrollo de herramientas computacionales para seleccionar y cambiar el criterio que debe ser utilizado por el sistema.

Dada las consideraciones mencionadas arriba, en [6] se introduce formalmente el concepto de Servicio de Razonamiento Rebatible basado en Preferencias. En particular, estos servicios buscan mejorar las habilidades de razonamiento de los sistemas argumentativos incorporando herramientas programables que le permitan al usuario configurar de manera dinámica los criterios de comparación con los que el sistema cuenta para resolver los conflictos entre argumentos. Para lograr esto, utilizan el sistema argumentativo llamado Defeasible Logic programming (DeLP) [4], el cual es un formalismo que se caracteriza por manejar las comparaciones entre argumentos de forma modular. Sin embargo, en DeLP una vez que se adopta un criterio (establecido en la configuración del sistema) acorde al dominio representado, no es posible cambiarlo dinámicamente por otro criterio. De hecho, en enfoques donde se han considerado varios criterios [7, 8], se utiliza una combinación fija y predefinida de los mismos.

Es importante mencionar que la utilidad práctica de estos servicios se puede ver efectivamente aplicada en diferentes escenarios, particularmente para los definidos en el desarrollo de sistemas recomendadores [9]. Esto se debe a que estos últimos usualmente operan creando un modelo de las preferencias de los usuarios con el objetivo de poder anticiparse a las necesidades y preferencias del usuario. Un Servicio de Razonamiento Rebatible basado en Preferencias puede ser implementado como un servidor de recomendación [10].

Un paso importante en el presente trabajo, es estudiar la aplicabilidad de los Servicios de Razonamiento Rebatible basados en Preferencias en el dominio de los sistemas de reconocimiento biométrico, específicamente en lo referente a la seguridad de los datos biométricos.

Se han investigado y desarrollado diversos sistemas de reconocimiento biométrico [11]. Sin embargo, los sistemas existentes podrían ser vulnerables a diversos ataques, como los de reproducción, y sufrir de intrusión en la privacidad de los usuarios; lo que dificulta seriamente su amplia aceptación en los usuarios finales.

La información biométrica utilizada en un sistema de reconocimiento es parte de la privacidad del usuario, que merece una protección especial. Si se filtra esta información privada, los atacantes pueden utilizarla maliciosamente, lo cual puede amenazar la seguridad de la información del usuario en otros sistemas y generar grandes pérdidas a los mismos [12]. Para hacer frente a este problema, los sistemas recomendadores constituyen una técnica

que ha demostrado ser muy prometedora. Estos sistemas pueden utilizarse para generar listas priorizadas de acciones de defensa [13], para detectar amenazas internas, para supervisar la seguridad de la red [14] y para acelerar otros análisis.

En base a lo expuesto arriba, un importante avance en este trabajo es mejorar las capacidades de razonamiento de este tipo particular de sistemas de recomendación. De este modo, se podría lograr un avance significativo para los sistemas recomendadores basados en argumentación dentro del área de Inteligencia Artificial y Ciencias de la Computación. A continuación, se repasarán los resultados obtenidos y el trabajo en progreso.

Líneas de Investigación y Desarrollo

La línea principal de investigación de este artículo toma como base el concepto de Servicio de Razonamiento Rebatible basado en Preferencia, y busca avanzar en la implementación de los diferentes mecanismos de manejo de preferencias que se han desarrollado en [6] para este tipo de servicio de razonamiento en particular. Por lo tanto, nuestra línea de investigación está motivada por tres ejes. Por un lado, en avanzar en la implementación de herramientas que permitan especificar de forma declarativa el criterio que el sistema debe utilizar. Luego, en desarrollar mecanismos para realizar una selección condicional basada en información contextual de uno de los criterios disponibles por el servicio. Por último, estudiar e implementar diferentes operadores de combinación criterios.

Servicio de Razonamiento Rebatible basado en Preferencias
Estos servicios proveen un razonamiento lógico rebatible con la capacidad de modificar las preferencias sobre la información procesada cambiando, a partir de cada consulta recibida, el criterio de preferencia utilizado. Tienen como objetivo permitir el desarrollo de servicios que respondan consultas de manera personalizada, y se puedan configurar automáticamente para que resuelvan el conflicto entre información contradictoria mediante el criterio que mejor se ajusta a un escenario particular.

Servicio de Razonamiento Rebatible basado en Preferencias Condicionales
Estos servicios extienden las capacidades de los Servicio de Razonamiento Rebatible basado en Preferencias, incorporando en las consultas una expresión formal de preferencia condicional que permite decidir cuál es el criterio de preferencia a ser utilizado en cada situación específica. De esta manera, el usuario puede guiar el proceso de razonamiento acorde a sus preferencias o necesidades.

Servicio de Razonamiento Rebatible basado en Preferencias Combinadas
En estos servicios de razonamiento, las consultas se caracterizan por incluir una expresión que permite combinar las especificaciones de varios criterios. El objetivo de este tipo de servicios es proveer un razonamiento basado en más de un criterio.

Resultados Esperados

El objetivo principal de este trabajo es abordar el estudio e implementación de herramientas computacionales concretas para seleccionar y cambiar el criterio de preferencia entre argumentos, utilizado en el sistema argumentativo DeLP. La idea general es brindar la posibilidad de programar, a través de expresiones condicionales y operadores específicos, qué criterio de preferencia se utilizará en cada situación particular. Cada uno de estos operadores representa una forma

diferente de combinar algebraicamente varios criterios. En este trabajo se estudiarán otros operadores, además de los propuestos en [6].

Las herramientas desarrolladas serán sometidas a análisis de correctitud, complejidad y eficiencia. Se analizarán los recursos computacionales de tiempo y espacio que cada solución exija, efectuando pruebas individuales e integradas de los mecanismos implementados.

Finalmente, se estudiará la aplicabilidad de las herramientas implementadas en el dominio de los sistemas de reconocimiento biométrico, específicamente en lo referente a la seguridad de los datos biométricos. Como resultado, se explorará la posibilidad de implementar un sistema concreto de recomendación para este dominio en particular.

Formación de Recursos Humanos

En la presente línea de investigación se enmarca el desarrollo de una tesis para la Maestría en Sistemas de la Información de la Universidad Nacional de Entre Ríos.

Referencias

1 C. Alvez, G. Etchart, S. Ruiz, E. Miranda and J. Aguirre, *Extensión de una base de datos Objeto-Relacional para el soporte de datos de iris*. XXIII Congreso Argentino de Ciencias de la Computación. Universidad Nacional de La Plata– Argentina, 2017.

2 S. Ruíz, G. Etchart, C. Alvez, E. Miranda, M. Benedetto and J. Aguirre, *Iris Information Management in Object-Relational Databases*. XXI Congreso Argentino de Ciencias de la Computación. Universidad Nacional del Noroeste de la Provincia de Buenos Aires, Junín - Argentina, 2015.

3 G. Etchart, L. Luna, R. Leal, M. Benedetto and C. Alvez, *Sistema adecuado a estándares de reconocimiento de personas mediante el iris*. CGIV - XIV Workshop de Investigadores en Ciencias de la Computación, Universidad Nacional de Entre Ríos, Concordia - Argentina, 2012.

4 A. Garcia, G.R. Simari. *Defeasible logic programming: An argumentative approach*. Theory and Practice of Logic Programming (TPLP) 4, 95-138, 2004.

5 I. Rahwan, G.R. Simari. *Argumentation in Artificial Intelligence*, 1st ed. Springer Publishing Company, Incorporated, 2009.

6 J.C. Teze. *Formalización y Generalización del Manejo de Preferencias en Servicios de Razonamiento Rebatible* (Tesis doctoral). Universidad Nacional del Sur, Bahía Blanca, Argentina, 2017.

7 C.A.D. Deagustini, S.E. Fulladoza Dalibón, S. Gottifredi, M.A. Falappa, C.I. Chesñevar, G.R. Simari, *Relational databases as a massive information source for defeasible argumentation*, Knowledge-Based Systems 51, pp. 91–109, 2013.

8 L. Godo, E. Marchioni, P. Pardo, *Extending a temporal defeasible argumentation framework with possibilistic weights.* Logics in Artificial Intelligence - 13th European Conference, JELIA 2012, Toulouse, France, September 26-28. Proceedings, pp. 242–254, 2012.

9 C.I. Chesñevar, A.G. Maguitman, G.R. Simari. *Recommender system technologies based on argumentation* 1. In: Emerging Artificial Intelligence Applications in Computer Engineering, pp. 50-73, 2007.

10 J. C. Teze, S. Gottifredi, A. J. García, G.R. Simari. *Improving argumentation-based recommender systems through context-adaptable selection criteria*, Expert Syst. Appl. 42 (21) pp. 8243–8258, 2015.
11 K. Delac and M. Grgic. *A survey of biometric recognition methods.* Proceedings. Elmar-2004. 46th International Symposium on Electronics in Marine, Zadar, Croatia, pp. 184-193, 2004.
12 Z. Rui, Z. Yan. *A Survey on Biometric Authentication: Towards Secure and Privacy*-Preserving Identification. IEEE Access, 7, 5994 – 6009, 2019.
13 K.B. Lyons. *A Recommender System in the Cyber Defense Domain*. Master's thesis no. AFIT-ENG-14-M-49, Air Force Institute of Technology Graduate School of Engineering and Management, Wright-Patterson Air Force Base, 2014.
14 T.A. Lewis. *An Artificial Neural Network-Based Decision Support System for Integrated Network Security*. Master's thesis no. AFIT-ENG-T-14-S-09, Air Force Institute of Technology Graduate School of Engineering and Management, Wright-Patterson Air Force Base, 2014.

Inteligencia Artificial de las Cosas

Azar Miguel Augusto[1], Tapia María Antonia[2], García Jorge Luis[3], Pérez Adrián Jesús Matías[4]
Facultad de Ingeniería - Universidad Nacional de Jujuy
Italo Palanca Nº 10 (+54 0388 4221587) – San Salvador de Jujuy - Jujuy
augusto.azar@gmail.com[1], maria.antonia.tapia@gmail.com[2], garciajorge4612@gmail.com[3], perezjadrian@gmail.com[4]

Resumen

El campo de la inteligencia artificial (AI) se aplica en diversos ámbitos en los cuales se registra la existencia de tareas repetitivas, alta presición, manejo de grandes volúmenes de información, riesgo de vida, extrema complejidad en la resolución de problemas, entre otros.

La tendencia actual indica que las diferentes técnicas de AI tales como los algoritmos genéticos, las redes neuronales, la lógica difusa, las hiperheurísticas y en general los sistemas de aprendizaje automático (Machine Learning) han escalado hacia el ámbito de Internet de las Cosas. El presente proyecto se encuentra enmarcado dentro de la investigación y el desarrollo de aplicaciones orientadas al uso de sensores, actuadores, interfaces y sistemas embebidos en IoT empleando algoritmos y técnicas de inteligencia artificial.

Esta investigación se enfoca por un lado en el diseño de interfaces hardware entre sensores y sistemas embebidos, codificación de los algoritmos de lectura de señales provistas por los sensores y transmisión inalámbrica a una base de datos para el almacenamiento de las señales obtenidas. Por otro lado, se concentra en el preprocesamiento de las señales registradas en la fase anterior, el análisis, codificación, pruebas y obtención de resultados mediante el diseño de algoritmos inteligentes.

Palabras clave: Artificial Intelligence, Embeded Systems, Sensors, Internet of Things, Smart City.

Contexto

Con la denominación “Diseño de Algoritmos Inteligentes Aplicados a Interfaces Humano Computador e Internet de las Cosas” [1] el proyecto dio inicio a principios de 2018. La investigación inicial consistió en el desarrollo de algoritmos inteligentes que permitan inferir conocimiento acerca de los sensores conectados a sistemas embebidos. En el mismo la I/D/I se realizó en base a una convocatoria de la Secretaría de Ciencia y Técnica de Facultad de Ingeniería (UNJu) a través del Expediente S-7710/17 y mediante las Resoluciones C.A.F.I. Nº 270/17 y C.A.F.I. Nº 271/17. La convocatoria se estableció para docentes incorporados al sistema de incentivos de la Secretaría de Políticas Universitarias, con categorias IV y V. También se enmarcó para dar cumplimiento a los compromisos de mejoramiento desde la Resolución

CONEAU Nº 1230/12 de acreditación de las Carreras de Ingeniería Informática y Licenciatura en Sistemas. La propuesta fue aprobada mediante Resolución C.A.F.I. Nº 661/17.

Si bien el proyecto se ejecutó durante 2018, se encuentra en trámite la solicitud de prórroga para 2019 debido a las dimensiones interdisciplinarias que el mismo contempla.

Introducción

Internet of Things fue el término usado en 1999 por Kevin Ashton [2] para hacer referencia a todo conjunto de dispositivos electrónicos dotados de sensores y conectados a Internet. Algunos autores definen IoT como una extensión de Internet hacia el ámbito físico [3], mientras que otros lo definen como el colectivo de sensores colocados en "cosas" y dentro de infraestructuras cibernéticas [4]. La relevancia de este concepto se ha intensificado con el transcurso del tiempo debido, entre otras razones, al creciente uso de sensores en smartphones y automóviles [5].

De este modo resulta ineludible admitir que IoT es un ámbito de gran relevancia en el que es necesario investigar. En [6], se afirma que en IoT los esfuerzos se concentran hacia tres categorías de investigación: i) Explorar las ideas de diseño. ii) Explorar sistemas a través del diseño. iii) Explorar componentes técnicos a través del diseño.

La evolución de IoT con el transcurso del tiempo se reflejó al dotar de "inteligencia" a los denominados end-devices (o end-nodes). Esto dio origen a los conceptos de Smart City, Smart Home, Smart Grid, Smart Lighting, Smart Parking, entre otros.

De esta forma las aplicaciones inteligentes en su mayoría ya no se implementan aisladamente, sino que operan conectadas a otros subsistemas o supra sistemas que hacen uso tanto de la información como de la meta información generada.

Por otro lado, en diversos foros [7, 8, 9, 10, 11, 12, 13 y 14] se hace referencia a la convergencia entre Inteligencia Artificial e Internet de las Cosas. Esta convergencia o combinación de técnicas, metodologías y estrategias de diseño en IoT junto a los algoritmos de inferencia inteligentes conforman el concepto de Inteligencia Artificial de las Cosas (AIoT).

Justificación del proyecto

Uno de los motivos por los cuales se continúa desarrollando el proyecto inicial reside en la creciente demanda de productos tanto en el campo de IoT como de AI. Este incremento se ve reflejado a través de estadísticas y tendencias las cuales sostienen que solo en 2016 el gasto en sistemas IoT ascendió a mas de 500 millones de dólares. Ese mismo gasto se espera que se eleve a 1.5 billones de dólares para 2020 [15]. Por otro lado, en una encuesta reciente realizada a CEOs de diferentes empresas se determinó que la importancia de las tecnologías del futuro para el éxito se concentra principalmente en dos campos, IoT y AI [16].

Es por esto que en el proyecto se tuvo en cuenta la incidencia del investigador y sus desarrollos en la sociedad; de modo que se optó por el análisis y desarrollo de un sistema de averías viales sin el cual el mantenimiento óptimo de calles y rutas queda supeditado al pedido formal del ciudadano.

Metodología

Diferentes metodologías son altamente adaptables a proyectos que involucran desarrollos no clásicos como los citados

(IoT y AI). Sin embargo, se ha optado por la incorporación de una metodología específicamente diseñada para el proceso de construcción de sistemas IoT.

Técnicamente se trata del paradigma Ignite [17]. Dicho enfoque, fue propuesto por Slama y colegas en 2015 y plantea subdividir el desarrollo y construcción del sistema en dos fases: Iniciación de la estrategia y Entrega de soluciones (Fig. 1).

La primera parte, 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 técnicamente para adoptar dichos proyectos. Esto implica básicamente la creación y gestión de un portafolio de proyectos de IoT.

La segunda fase, orientada mas al aspecto estrictamente tecnológico, direcciona a los gerentes de producto y jefes de proyecto en la planificación y ejecución. Tal como se observa en la Figura 1, se compone de tres fases las cuales son: Plan, Build y Run [18].

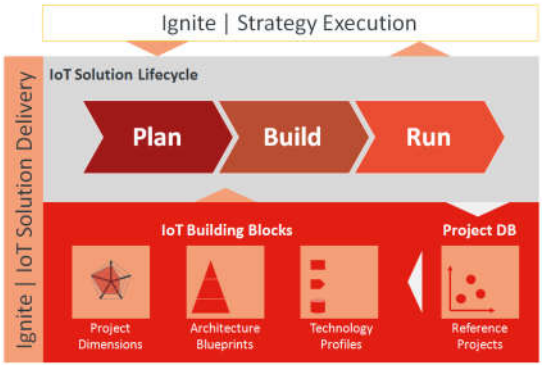


Figura 1 – Etapas de la fase Entrega de la Solución (metodología Ignite).

En el proyecto el esfuerzo esta centrado en la segunda fase dado que especifica los pasos y detalles técnicos para el desarrollo de sistemas que involucran conectividad de sensores.

Líneas de Investigación, Desarrollo e Innovación

Una de las líneas de investigación que se inició y se encuentra en proceso es la de Smart Cities haciendo énfasis en Internet de las Cosas Móviles (IoMT) [19]. Mas específicamente se trabajó sobre una primera fase para el desarrollo de un sistema de detección de averías viales (baches) mediante la lectura de acelerómetro incorporado en smartphones [20]. Este primer acercamiento permitió determinar la posible factibilidad de uso del acelerómetro como sensor a conectar a un sistema embebido (en fases posteriores). Si bien, luego de los experimentos iniciales el acelerómetro resultó una opción satisfactoria, las otras alternativas de elección de un transductor de medición de las averías viales que se manejaron desde el inicio y aun se encuentran en proceso de estudio son las siguientes:

- Sensor cerámico piezoeléctrico [21]
- Sensor de vibración (SW18020P) [22]
- Módulo acelerómetro(MMA7361) [23]
- Giroscopio (MPU6050) [24]

En cuanto a las placas de desarrollo se dispone de 4 computadoras industriales EDU-CIAA, 4 Raspberry PI (diferentes versiones) y sensores varios. Actualmente el proyecto se encuentra en etapa de pruebas mediante Raspberry PI en combinación con el módulo acelerómetro y módulo GPS.

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.

El cúmulo de conocimientos adquiridos hasta el momento presenta la oportunidad de iniciar el dictado cursos-talleres que se tenían previstos al comienzo del proyecto. En el proceso de adquisición cognitiva se obtuvieron conceptos tanto en investigación básica como aplicada.

Formación de Recursos Humanos

Actualmente el grupo de investigación esta formado por un director, una egresada de Licenciatura en Sistemas, un alumno de Ingeniería Informática y un alumno de Licenciatura en Sistemas.

Referencias

[1] Azar M.A., Medrano J.F., Tapia M.A., Carlos F.J., Villafañe J.P., Mamaní A.I., Aparicio F., Santos W. B., Estrada B.R. y Molloja J. Diseño de Algoritmos Inteligentes Aplicados a Interfaces Humano Computador e Internet de las Cosas. Workshop de Investigadores de Ciencias de la Computación (WICC). pp 17-21. 2018.

[2] Ashton K. That 'Internet of Things' Thing: In the real world, things matter more than ideas. RFiD Journal. RFID. 2009.

[3] Tsirmpas C., Anastasiou A., Bountris P. y Koutsouris D. A New Method for Profile Generation in an Internet of Things Environment: An Application in Ambient-Assisted Living. IEEE Internet of Things Journal. Volume: 2, Issue: 6. Page(s): 471 – 478. ISSN: 2327-4662. 2015.

[4] Nahrstedt K., Li H., Nguyen P., Chang S. y Vu L. Internet of mobile things: Mobility-driven challenges, designs and implementations. In Proceedings - 2016 IEEE 1st International Conference on Internet-of-Things Design and Implementation, IoTDI 2016 (pp. 25-36). [7471348] Institute of Electrical and Electronics Engineers Inc. DOI: 10.1109/IoTDI.2015.41. 2016.

[5] Wearable Sensors Market by Type - Global Forecast to 2022. Online. http://www.marketsandmarkets.com/Market-Reports/wearable-sensor -market-158101489.html. Consultado el 03 de febrero de 2019.

[6] 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.

[7] ITU Forum on Artificial Intelligence, Internet of Things and Smart Cities. https://www.itu.int/en/ITU-T/ Workshops-and-Seminars/201812/ Pages/ Programme.aspx. Consultado el 13 de febrero de 2019.

[8] Global Artificial Intelligence of Things (AIoT) Solutions Market Report 2018. https://www.prnewswire.com /news

-releases/global-artificial-intelligence-of-things-aiot-solutions-market-report-2018-global-ai-in-embedded-iot-devices-market-will-approach-26-2b-usd-by-2023--300794268.html. Consultado el 02 de marzo de 2019.

[9] How the Artificial Intelligence of Things (AIoT) Boosts the Value of Your IoT Data. https://www.arcweb.com/blog/how-artificial-intelligence-things-aiot-boosts-value-your-iot-data. Consultado el 09 de marzo de 2019.

[10] Artificial Intelligence of Things (AIoT). https://internetofthingsagenda.techtarget.com/definition/Artificial-Intelligence-of-Things-AioT. Consultado el 22 de febrero de 2019.

[11] Toward the artificial intelligence of things. https://www.infoworld.com/article /3269325/toward-the-artificial-intelligence-of-things.html. Consultado el 10 de febrero de 2019.

[12] The Artificial Intelligence of Things. White paper SAS Institute Inc. 2018.

[13] Introducing the Artificial Intelligence of Things. http://www.b2b.com /introducing- the- artificial - intelligence-of-things. Consultado el 18 de febrero de 2019.

[14] Guo Y., Liu H. Y Chai Y. The Embedding Convergence Of Smart Cities And Tourism Internet Of Things In China: An Advance Perspective. Advances in Hospitality and Tourism Research (AHTR), 2(1): 54-69, An International Journal of Akdeniz University Tourism Faculty ISSN: 2147-910. 2014.

[15] Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. https://www.gartner.com/en/newsroom/press-releases/2017-02-07-gartner-says- 8-billion-connected-things-will-be-in- use-in-2017-up-31-percent-from-2016. Consultado el 05 de febrero de 2019.

[16] Top 10 Tech Trends from CES 2019. https://www.emarketer.com/content /top-10-tech-trends-from-ces-2019. Consultado el 02 de marzo de 2019.

[17] Ignite | IoT Methodology. http://enterprise-iot.org/book/enterprise-iot/ part -ii-igniteiot- methodology/. Consultado el 11 de noviembre de 2018.

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

[19] Nahrstedt K. Internet of mobile things: challenges and opportunities. 10.1145/2628071.2635931. 2014.

[20] Azar M.A. y Tapia M.A. Detección de Averías Viales Mediante IoMT Aplicada a Smart Cities. Congreso Argentino de Ciencias de la Computacion CACIC. ISBN: 978-950-658-472-6. Tandil. 2018.

[21] Piezoelectric Ceramic Sensors. Datasheet. http://www.farnell.com/datasheets /43406.pdf. Consultado el 04 de febrero de 2019.

[22] Vibration Sensor switch. Datasheet. https://e-radionica.com/productdata/ SW -18020 .pdf. Consultado el 16 de febrero de 2019.

[23] MMA7361 3-Axis Accelerometer Module. Datasheet. https://eecs.oregonstate.edu/education/docs/accelerometer /MMA7361_module.pdf. Consultado el 03 de febrero de 2019.

[24] MPU-6000 and MPU-6050 Product Specification Revision 3.4. https://store.invensense.com/datasheets /invensense/MPU-6050_DataSheet_ V3% 204 .pdf. Consultado el 02 de marzo de 2019.

Minería de Opiniones: Análisis de Sentimientos en una Red Social

Alejandra Cardoso [1], Lorena Talame [1], Matias Amor [1], Carlos Neil [1,2]

Grupo de Análisis de Datos /Facultad de Ingeniería e IESIING
[1] Universidad Católica de Salta
Campo Castañares s/n, 4400 Salta, (0387) 426 8536.
{acardoso, mltalame}@ucasal.edu.ar, matiasnicolasamor@gmail.com
[2] Universidad Abierta Interamericana
Carlos.Neil@uai.edu.ar

RESUMEN

El Grupo de Análisis de Datos de la Facultad de Ingeniería y del Instituto de Estudios Interdisciplinares de Ingeniería (IESIING) de la Universidad Católica de Salta, trabaja en proyectos de investigación relacionados a técnicas y aplicaciones de minería de datos y minería de textos. Entre las áreas investigadas se incluyen la categorización de documentos de texto, búsqueda semántica, extracción de entidades con nombre, generación de resúmenes y búsqueda automática de respuestas. Con el proyecto actual se añade un nuevo aspecto en la línea de investigación: el análisis de sentimientos también llamado minería de opiniones. En el tratamiento de opiniones, muchas investigaciones se centraron en el reconocimiento de polaridad en textos. Sin embargo, pocas realizaron una clasificación más exhaustiva y, mucho menos, en lenguaje español. El objetivo del proyecto es clasificar textos cortos (opiniones) en seis sentimientos: miedo, ira, asco, sorpresa, tristeza y felicidad. El proyecto se encuentra en desarrollo. En la primera etapa se recopilaron mensajes de la red social Twitter, muchos de los cuales formarán el corpus de análisis. En la etapa actual se están evaluando diversos enfoques para clasificación y detección de sentimientos en los mismos.

***Palabras clave*:** análisis de sentimientos, minería de textos, Twitter.

CONTEXTO

Este proyecto de investigación continúa la línea de investigación que el Grupo de Análisis de Datos viene desarrollando en minería de textos, financiados por el Consejo de Investigaciones de la Universidad Católica de Salta. El primer trabajo de investigación fue sobre el problema de clasificación o categorización automática de documentos [1]. El segundo versó sobre el reconocimiento de entidades con nombre [2] y la extracción automática de resúmenes de documentos de texto [3]. La plataforma para experimentar con las técnicas descritas fue un buscador semántico en el corpus de más de 8000 documentos conteniendo nueve años de resoluciones rectorales de la Universidad Católica de Salta en distintos formatos. En el tercero, se exploró la búsqueda automática de respuestas a preguntas de usuarios en lenguaje natural en el mismo corpus [4] y en textos extraídos de la web [5]. En esta oportunidad,

el enfoque es el análisis y clasificación de sentimientos de las opiniones de usuarios emitidos en una red social.

1. INTRODUCCION

Debido a la enorme cantidad de documentos disponibles en forma digital y la también creciente necesidad de organizarlos y aprovechar el conocimiento contenido en ellos, en los últimos años, ha aumentado el interés en las técnicas de análisis de información textual. La minería de textos es el proceso de extraer información y conocimiento interesante y no trivial de texto no estructurado. Es un campo relativamente reciente con un gran valor comercial que se nutre de las áreas de recuperación de la información, minería de datos, aprendizaje automático, estadística y procesamiento del lenguaje natural.

La minería de textos incluye una serie de tecnologías: extracción de la información, seguimiento de temas (topic tracking), generación automática de resúmenes de textos (sumarización), categorización, agrupamiento (clustering), vinculación entre conceptos, visualización de la información, respuesta automática de preguntas y análisis de sentimientos. El presente trabajo se centra en la última, también denominada minería de opiniones.

La información puede dividirse en dos grandes grupos, la información objetiva, que representa un hecho, algo comprobable, por ejemplo "*Salta es una provincia del norte argentino*", y la información subjetiva que puede entenderse como la interpretación de una realidad, por ejemplo "*Salta es la provincia más linda del norte argentino*". La minería de opiniones consiste en el análisis de la información subjetiva (opiniones). Un caso particular son las opiniones que generan usuarios de redes sociales, blogs, portales de noticias, foros, etc. que ayudan a revelar información importante sobre un tema específico apoyando, por ejemplo, a campos como la inteligencia de negocios y jugando un papel importante en la toma de decisiones. Una forma de monitorear la opinión de los usuarios sobre determinado producto o tema es proponer encuestas o resaltar comentarios de otros usuarios, por ejemplo otorgando puntaje. Si bien esto puede ser válido, la opinión textual de los usuarios sobre un tema en cuestión se la puede clasificar como positiva o negativa y/o realizar una clasificación más avanzada identificando sentimientos como por ejemplo, tristeza, miedo, sorpresa, alegría, etc. La red social Twitter se ha convertido en una excelente herramienta para conocer en tiempo real las opiniones que los usuarios expresan sobre una gran variedad de temas. Twitter permite identificar estos temas a través de los denominados hashtag o etiquetas, que se caracterizan por comenzar con el carácter # y a continuación una cadena de una o varias palabras concatenadas.

La mayoría de los enfoques actuales se basan principalmente en dos tipos de tareas: detección de la polaridad de la opinión y análisis de sentimientos basado en características.

1) Detección de la polaridad: consiste en determinar si una opinión es positiva o negativa. Algunos trabajos extienden la polaridad básica a un rango de valoración, por ejemplo, en [6] realizaron experimentos con un corpus en español de críticas de cine obtenidas de un sitio web sobre novedades, cartelera y opiniones sobre cine. Las críticas fueron puntuadas en un rango de 1 a 5,

significando el 1 una película muy mala, y el 5 una película muy buena. Las películas puntuadas con 3 se catalogaron como neutras, o que el crítico (usuario de la web) no las consideraba ni malas ni buenas.

Para obtener la polaridad, existen dos métodos más populares: el aprendizaje computacional y el basado en diccionarios léxicos.

- El aprendizaje computacional analiza la información automáticamente de forma supervisada, basándose en conjuntos de entrenamiento que son utilizados para catalogar al resto de las opiniones, realizando pruebas y luego validándolas. Un ejemplo de aplicación de algoritmos de aprendizaje supervisado se encuentra en [7].

- El método de diccionarios léxicos se basa en una lista de palabras con un determinado peso y/o categoría emocional. Estos diccionarios presentan principalmente adjetivos, que son los que aportan mayor información al momento de analizar los sentimientos, aunque también incluye verbos, adverbios y sustantivos. En [8] se utilizó un diccionario de palabras positivas y negativas con puntaje para determinar la polaridad de mensajes de Twitter durante un acto electoral.

2) Análisis del sentimiento basado en características: consiste en determinar las distintas características o entidades del producto tratadas en la opinión escrita por el usuario, y para cada una de esas características mencionadas en la opinión, ser capaces de extraer una polaridad. En [9] se recopilaron opiniones de un sitio gastronómico y se definieron las entidades comida, ambiente y servicio, así de cada comentario se identificó la opinión (positiva, negativa, neutra) sobre cada entidad.

Algunas investigaciones se centraron en analizar opiniones y detectar algún sentimiento en particular. En [10] se clasificó una serie de palabras según seis sentimientos: alegría, enojo, miedo, tristeza, sorpresa y repulsión formando un diccionario de emociones. El trabajo de [11] propone un modelo conceptual para la detección de mensajes violentos o peligrosos. El análisis realizado por [12] se enfocó en diferenciar distintos sentimientos (me encanta, sublime, patético, lamentable, etc.) manifestados por los usuarios.

La propuesta del presente proyecto es llegar un poco más allá de la extracción de polaridad en mensajes de textos, de tal forma de detectar el sentimiento que se expresa en los mismos. Para ello se están analizando las distintas alternativas mencionadas para la clasificación. Así, servirá como puntapié inicial para otro tipo de tareas que posibilite, por ejemplo, eliminar textos agresivos, tan frecuentes en una red social.

2. LINEA DE INVESTIGACION Y DESARROLLO

Este proyecto de investigación propone detectar y clasificar sentimientos expresados en textos, en particular, opiniones textuales emitidas en una red social. El proyecto se desarrolla con las siguientes tareas:

- Revisión de la literatura relevante al problema de minería de opiniones y sentimientos.
- Exploración y evaluación de las diferentes técnicas de recopilación de mensajes
- Recopilación de mensajes y creación de un corpus de opiniones
- Exploración y evaluación de las técnicas de clasificación de textos.

Hasta el momento se logró capturar más de 60000 tweets utilizando la API[1] de Twitter, los cuales fueron recopilados, diariamente durante tres meses, a partir de una serie de hashtags sobre temas de actualidad de nuestro país. Los mensajes se almacenaron en una base de datos NoSQL. Muchos de estos tweets serán descartados por contener solo imágenes, íconos o textos con poca información para el análisis.

Actualmente, el grupo de investigación se encuentra en la etapa de limpieza y preparación de los tweets, y evaluando distintos algoritmos de clasificación de textos.

3. RESULTADOS OBTENIDOS/ESPERADOS

El proyecto tiene como objetivo general, analizar mensajes de textos generados en la red social Twitter e identificar los sentimientos que se expresen en ellos.

Una de las primeras etapas consistió en la captura de tweets, para lo cual se evaluó distintas formas de recopilación.

Los próximos objetivos a alcanzar son:

- Realizar la limpieza y preparación de los tweets descartando aquellos que no posean suficiente información textual para el análisis.
- Evaluar las formas de etiquetado de los mensajes que formarán el conjunto de entrenamiento para los algoritmos.
- Evaluar y comparar algoritmos de clasificación de textos
- Seleccionar los algoritmos que mejor clasifiquen las opiniones

Se espera que esta línea de investigación amplíe los conocimientos sobre las diferentes técnicas de minería de textos y procesamiento de lenguaje natural.

Por otro lado, se espera que este proyecto anime el interés por la investigación y por esta temática a los alumnos de nuestra Facultad.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está integrado por tres docentes de la carrera de Ingeniería en Informática y dos alumnos de la carrera, una de ellas se encuentra realizando su proyecto de grado en la temática de esta investigación.

REFERENCIAS

[1] A. Pérez Abelleira y A. Cardoso, «Categorización automática de documentos,» de *Simposio Argentino de Inteligencia Artificial, 40 Jornadas Argentinas de Informática (JAIIO)*, Córdoba, 2011.

[2] A. Pérez Abelleira y A. Cardoso, «Técnicas de extracción de entidades con nombre,» de *Simposio Argentino de Inteligencia Artificial, 42 Jornadas Argentinas de Informática (JAIIO)*, Córdoba, 2013.

[3] A. Cardoso y A. Pérez Abelleira, «Generación automática de resúmenes,» de *Congreso Nacional de Ingeniería Informática/Sistemas de Información (ConNaIISI)*, Córdoba, 2013.

[4] A. Cardoso, A. Bini y A. Pérez Abelleira, «Una Arquitectura de un Sistema de Búsqueda de Respuestas,» de *2° Congreso Nacional de Ingeniería Informática/ Sistemas de Información, CoNaIISI*, San Luis, 2014.

[5] A. Cardoso, A. Pérez Abelleira y E. Notario, «Búsqueda de respuestas como aplicación del problema de extracción de relaciones,» de *4º Congreso Nacional de Ingeniería Informática/Sistemas de Información (ConNaIISI)*, Salta, 2016.

[6] E. Cámara, M. Valdivia, J. Ortega y A. Ureña Lopez, «Técnicas de clasificación de opiniones aplicadas a un corpus en español,» *Procesamiento del Lenguaje Natural,* nº 47, pp. 163-170, 2011.

[7] T. Baviera, «Técnicas para el análisis del

[1] https://developer.twitter.com/

sentimiento en Twitter: Aprendizaje Automático Supervisado y SentiStrength,» *Digitos. Revista de comunicación digital,* vol. 1, nº 3, pp. 33-50, 2017.

[8] L. Montesinos García, «Análisis de sentimientos y predicción de eventos en Twitter,» Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile, Santiago de Chile, 2014.

[9] L. Dubiau, *Procesamiento de Lenguaje Natural en Sistemas de Análisis de Sentimientos,* Buenos Aires: Facultad de Ingenieria, Universidad de Buenos Aires, 2013.

[10] I. Diaz Rangel, G. Sidorov y S. Suarez Guerra, «Creación y evaluación de un diccionario marcado con emociones y ponderado para el español,» *Onomazein. Revista semestral de lingüística, filología y traducción,* pp. 31-46, 2014.

[11] J. C. Calloni, A. Bianciotti, S. Páez, E. Scarello, L. Banchio, M. Mulassano, J. Saldarini, F. Francia, F. Degiovanni, L. Scharff y J. C. Cuevas, «Modelo de análisis de sentimientos con algoritmos de aprendizajes para detectar actitudes peligrosas o violentas de los usuarios en redes sociales,» de *IV Congreso Nacional de Ingeniería Informática y Sistemas de Información (CONAIISI)*, Salta, 2016.

[12] SM Reputation Metrics, «SM Reputation Metrics,» 2015. [En línea]. Available: https://smreputationmetrics.wordpress.com/2015/10/19/el-debate-albertvspablo-analisis-de-sentimiento-twitter/.

Modelo de Redes Neuronales Convolucionales Profundas para la Clasificación de Lesiones en Ecografías Mamarias

Hugo Chanampe[1], Silvana Aciar[2], Martin de la Vega[3], José Luis Molinari Sotomayor, German Carrascosa, Alejandro Lorefice
Centro de Investigación y Desarrollo Informático (CIDI)[13]
Universidad Nacional de La Rioja[13]
Universidad Nacional de San Juan[2]
[1]hugochanampe@unlar.edu.ar
[2]saciar@unsj-cuim.edu.ar
[3]martindelavega@unlar.edu.ar

Resumen

El cáncer de mama es una de las principales causas de muerte en mujeres mayores en Argentina. Nuestro estudio apunta a utilizar técnicas Deep Learning para clasificar distintos tipos de lesiones en ecografías mamarias. Se muestra el actual estado del arte de las arquitecturas de Deep Learning en la clasificación de imágenes y sus distintas estrategias de entrenamiento. Usando Redes Neuronales Convolucionales Profundas se desarrolló un modelo que permite clasificar diferentes tipos de lesiones. El modelo logro alcanzar una precisión del 62%, sobre la base datos mini-MIAS

1. Contexto

Este artículo se encuentra enmarcado dentro del proyecto de investigación *"Deep Learning para el Reconocimiento de Patrones"*, aprobado según resolución *CICYT 056/17* perteneciente al *Departamento Académico de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de La Rioja*. Su ejecución se está llevando a cabo en el *Centro de Investigación y Desarrollo Informático (CIDI)* perteneciente a este mismo departamento.

2. Introducción

Hoy en día, las técnicas de *Deep Learning* (DL) o aprendizaje profundo vienen logrando exitosos resultados en el procesamiento de imágenes naturales, reconocimiento de objetos dentro de las mismas y clasificación de imágenes. Todo esto es posible gracias a las grandes capacidades de cómputo actuales y también gracias a la convivencia de las potentes GPU de la actualidad.

Uno de los métodos de DL que se emplea para el análisis de imágenes son las *Redes Neuronales Convolucionales Profundas* (DCNN) [1], que puede descubrir las características en las imágenes, como detección de bordes y esquinas hasta llegar a detectar características más complejas, a través del entrenamiento con grandes conjuntos de datos. Esto también se está empleando en el análisis de imágenes médicas con prometedores resultados sobre varias de sus aplicaciones, segmentación de órganos, detección y clasificación de distintos tipos lesiones en los órganos, etc. [2]

El *Cáncer de Mama* (CDM) es una de las primeras causas de muerte en Argentina, en mujeres de entre 20 y 59 años de edad y la segunda para mujeres mayores de 59 años [3]. El diagnóstico y el tratamiento de esta enfermedad en etapas tempranas es vital para prevenir el progreso de la enfermedad y reducir los índices de mortalidad [4]. La revisión de mama se realiza a través del estudio de un tipo de imagen, como lo es *Ecografía Mamaria* (EM), que es el examen básico para el CDM.

La implementación de los métodos de DL, como las CNN puede ayudar a mejorar la clasificación de distintos tipos de lesiones dentro de las EM.

El propósito de este trabajo es desarrollar un modelo que use DL para clasificar los distintos tipos de lesiones en ecografías mamarias.

Este articulo está estructurado de la siguiente forma: En la sección 3 se presentan algunos conceptos de DL y una revisión de las diferentes arquitecturas de DL para el análisis de imágenes y diferentes estrategias de entrenamiento. En la sección 4 se presenta el modelo propuesto en este artículo. En la sección 5 se presentan los resultados obtenidos. En la sección 6 conclusiones y

trabajos futuros. Por último, en la sección 7 formación de recursos humanos y estructura del equipo de trabajo del proyecto de investigación.

3. Deep Learning o Aprendizaje Profundo

Los métodos de DL buscan un conjunto de transformación directamente desde los datos mismos. Esta mirada está logrando excelentes resultados, particularmente en los problemas que pertenecen al campo de *Visión de las Computadoras* (VC), como la clasificación de escenas naturales y la detección de objetos [5]. Los modelos de DL están siendo adaptados también a diferentes tareas médicas, como ayuda en el diagnóstico temprano del Alzheimer [6], segmentación automática en lesiones provocadas por la Esclerosis Múltiple [7], detección de tumores en ecografías mamarias (EM) [8] y muchos más. Sin embargo, únicamente unos pocos trabajos tienen explorados los métodos de DL en dirección de la clasificación automática de lesiones identificadas en EM [9].

Los métodos tradicionales de *Aprendizaje Automático* (ML) como las máquinas de soporte vectorial, el método del vecino más próximo y el análisis de discriminación lineal fueron usados como clasificadores para una gran variedad de desarrollos para el diagnóstico temprano del cáncer de mama [10] [11]. Sin embargo, tienen una capacidad limitada para procesar los datos en su forma original.

DL, sin olvidar que es un subcampo de ML, tiene métodos como lo son las arquitecturas de DCNN, que funcionan aprendiendo características relevantes directamente desde las grandes bases de datos (datos en bruto). Un aspecto clave de DL es que el aprendizaje de las características de los datos, no fue diseñado por un ingeniero experto. Estas fueron aprendidas desde los datos usando un procedimiento de aprendizaje general.

3.1. Antecedentes de las Redes Neuronales Convolucionales

Las DCNN se han convertido en una de las técnicas más populares dentro del campo de ML, que es muy usada en el reconocimiento de imágenes. Las capacidad de las DCNNs puede ser controlada variando su profundidad y amplitud, también estas redes pueden hacer generalmente suposiciones correctas [12].

La clasificación de imágenes o clasificación de exámenes fue uno de los primeros campos donde DL hizo una mayor contribución en el análisis de imágenes médicas. En un examen de clasificación, por lo general se tiene una o varias imágenes como entrada y una sola variable de salida, el diagnostico (ej.: presencia de enfermedad o ausencia). En [12] se representó un punto de inflexión para el reconocimiento de objetos a gran escala cuando se pudo entrenar una de las primeras DCNN sobre un gran base de datos como ImageNet.

Las DCNN requieren una gran cantidad de datos de entrenamiento etiquetados, un aspecto critico que puede ser un obstáculo difícil de sortear en el dominio de las imágenes médicas. Sin embargo, también el entrenamiento de las DCNN involucra muchos recursos de procesamiento y memoria, esto a veces podría consumir bastante tiempo. A veces el entrenamiento de este tipo de arquitecturas es complicado también por los problemas sobreajuste (overfitting) y bajo ajuste (underfitting) cuando hay datasets muy pequeños para el entrenamiento y cuando hay poco tiempo de entrenamiento [13]. Por eso el entrenamiento desde de este tipo de arquitecturas puede ser tedioso y consumir mucho tiempo, exigiendo una gran experticia y paciencia [9] [14].

Las DCNN es la técnica de DL más exitosa aplicada a la compresión de las imágenes [15]. Estas tuvieron su primera aparición exitosa en el mundo real en la aplicación de LeNet [16], para el reconocimiento de dígitos escritos a mano. A pesar de este éxito inicial no cobro impulso hasta que no se desarrollaron varias técnicas nuevas para el entrenamiento eficiente de las redes profundas. La clave fue el aporte realizado por [12] a fines del 2012 en la competencia de ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). La DCNN que presentaron en esa oportunidad, llamada AlexNet, gano la competición por un amplio margen. En el 2014 el Visual Geometry Group a través de su modelo propuesto (VGG-Net), cuyo principal aporte fue una evaluación a fondo de las redes de profundidad creciente usando una arquitectura muy pequeña de 3x3 filtros convulsionados (convolutional filters), que alcanzaron una significante mejora, esto se logró configurando profundidades de entre 16 y 19 capas de pesos (weigh layers) [17]. Cabe destacar que este modelo obtuvo el segundo lugar en la competencia de ILSVRC. También en el ILSVRC 2014 También en el ILSVRC 2014 la arquitectura ganadora propuso una Deep Convolutional Neural Network, que fue llamada GoogLeNet, que tenía una red con 22 capas de profundidad, que contiene dos convolutional layers, tres pooling layers y nueve inception layers, pero reduciendo el número de parámetros a 4 millones. Cada inception layer consta de seis convolutional layers y una pooling layer. Pero su marca principal, además de mejorar el estado del arte en la clasificación y detección en la competencia, fue la mejora en la utilización de los recursos computacionales dentro de la red [18]. En el 2015 se desarrolló Inception-V2, la primera versión de Inception fue GoogleNet anteriormente mencionada. Los cambios que se introdujeron en esta nueva versión fue la implementación del método llamado Batch Normalization (BN) que permite usar índices de aprendizajes mucho más altos, también actúa como un regularizador en las capas de entrada y hasta puede evitar uso de la técnica de Dropout

en la fase entrenamiento [19]. Usando el método BN se alcanzó la misma precisión que su antecesor, pero en menos pasos de entrenamiento de la red. Usando el modelo de Inception-V2 se logró un índice de error de validación 4.9 % sobre ImageNet.

Tabla 1 - Arquitecturas de Redes Neuronales Convolucionales Profundas

Algoritmos	Nº de Capas	Índice de Error	Posición en Competencia ILSVRC	Año
AlexNet	8	15.3 %	Primer Lugar	2012
VGGNet	19	7.3 %	Segundo Lugar	2014
GoogleNet	22	6.67 %	Primer Lugar	2014
ResNet	152	3.57 %	Primer Lugar	2015

Luego en la tercera versión, llamada Inception-V3, se mejoró la arquitectura agregando la idea de factorización que consiste en disminuir el número de parámetros en cada módulo inception, logrando reducir el costo computacional como se muestra en el trabajo [20]. Donde se presentó el modelo Inception-V3, que obtuvo mejores resultados sobre la misma base de datos con un índice de error sobre el set de validación de 3.58%. En el mismo año surge el problema de la degradación en las arquitecturas de redes neuronales profundas, que es cuando se aumenta la profundidad de una red neuronal (incremento en la cantidad de capas). Pero esto conlleva un perjuicio en la tasa de error del modelo, aumentando considerablemente como se reportó en [21] [22]. Aquí es donde aparecen las Deep Residual Nets que pueden fácilmente ganar precisión incrementando la profundidad en su arquitectura. La arquitectura que gano en 2015 la competencia ILSVRC fue ResNet-34. Esta arquitectura es muy similar a la VGG-Net, excepto porque a cada filtro de 3x3 se le agrega una conexión de acceso directo, lo que la convierte a la red en una versión residual de la anterior. Esta red alcanzo una considerable mejora en la tasa de error en el entrenamiento, esto indica que el problema de la degradación está bien direccionado, permitiéndole ganar en precisión y profundidad en la red [23]. En la *Tabla 1,* se muestra la progresión de las arquitecturas que fueron más transcendentales en estos últimos años, por los resultados alcanzados en competencia ImageNet, destacándose la disminución de los índices de error.

Hoy por hoy, las DCNNs están siendo más ampliamente usadas en el análisis de imágenes médicas

3.2. Estrategias de Entrenamiento en las Redes Neuronales Convolucionales

3.2.1. Pre-Training y Fine-Tuning

Una prometedora alternativa al entrenamiento desde cero de las DCNN es la transferencia de aprendizaje, que es esencialmente el uso de redes pre-entrenadas desde campos de aplicación diferentes, por lo general se lleva a cabo sobre imágenes naturales. Los modelos pre-entrenados están siendo aplicados de manera exitosa en varias tareas dentro del campo de visión de las computadoras (VC), que puede ser como un generador de características o como una base para el aprendizaje mediante transferencia [24] [25]. Se tiene de esta manera dos estrategias aplicadas a la trasferencia de aprendizaje que se pueden identificar: (1) uso de las redes pre-entrenadas (pre-training) solamente como extractor de características y (2) hacer pequeños ajustes (fine-tuning), sobre los pesos de una red pre-entrenada, para lograr una mejor performance sobre los datos a procesar. Pre-training significa inicializar las redes con parámetros pre entrenados en lugar de establecer parámetros aleatoriamente. Es bastante popular en los modelos basados en DCNN, debido a que da la ventaja de permitir acelerar el proceso de aprendizaje y mejorar la capacidad de generalización. Como AlexNet, logro excelentes resultados [12], numerosos enfoques eligen a este modelo entrenado sobre ImageNet 2012 como su modelo profundo de referencia [26] [27] y utilizan fine-tuning para el ajuste de los parámetros, en concordancia con la tarea específica que tienen como objetivo.

3.2.2. Data Augmentation

El aumento de datos es el método más sencillo para reducir el sobreajuste cuando trabajamos con imágenes como datos. Consiste en utilizar las mismas imágenes del dataset que tenemos y conservando sus etiquetas se aplican diferentes transformaciones en las imágenes para generar datos adicionales sin introducir costos adicionales de etiquetado. Por ejemplo en el modelo AlexNet [12], tiene dos formas distintas de aumento de datos: la primera consiste en generar desplazamientos horizontales de las imágenes y reflejos horizontales, y la segunda forma consiste en alterar la intensidad de los canales de RGB en las imágenes de entrenamiento [13]. También se tomó AlexNet como base y se agregó transformaciones adicionales que mejora la cualidad invariable del desplazamiento y el color así extendiendo con un montón de imágenes con pixeles extras y agregando manipulaciones de color adicionales [28].

3.2.3. DropOut

La implementación de este procedimiento de entrenamiento consiste en desconectar un porcentaje de las neuronas en cada iteración del entrenamiento. El termino Dropout [29] se refiere a sacar de una red neuronal, unidades ocultas o visibles. También esta técnica es utilizada para evitar el sobreajuste, que es cuando tiene un conjunto de datos reducidos para entrenar nuestro modelo. Esto ocasiona muchas veces que se produzca un sobreajuste sobre este set de datos por parte del modelo.

A continuación, en la *Tabla 2*, se muestra las diferentes estrategias de entrenamiento de las DCNN y su principal característica.

Tabla 2 - Estrategias de Entrenamiento de las Redes Neuronales Convolucionales Profundas

Técnicas	Características
Pre-training	Uso de las de redes solamente como extractor de características
Fine-Tuning	Se hacen pequeños ajustes sobre los pesos de una red.
Data Augmentation	Esta estrategia de entrenamiento se utiliza para agrandar el conjunto de datos de entrenamiento de manera artificial sobre imágenes.
DropOut	Permite reducir el overfitting, desconectando unidades visible como ocultas de una red neuronal.

4. Modelo Propuesto

En esta sección se propone el desarrollo de un modelo que empleará como datos entrada las EM y aplicará una o varias arquitecturas de DCNN, vistas en las secciones anteriores. Permitiendo obtener una salida la cual será discriminar los distintos tipos de lesiones que se detectaron en las imágenes de entrada del modelo.

4.1. Base de Datos de Mamografías

La base de datos que se utilizará para el análisis de distintos tipos de lesiones será mini-MIAS [30]. Los tipos de anormalidades que tiene este conjunto de dato son: 1) calcificación (CALC), 2) masas bien definidas (CIRC), 3) masas espiculadas (SPIC), 4) otras masas mal definidas (MISC), 5) Distorsión Arquitectónica (ARCH), 6) Asimetría (ASYM) y 7) Normal (NORM).

La base de datos contiene 322 imágenes que tienen una resolución de 1024x1024. Este base de datos es bastante pequeña en el contexto de DL y VC, pero es ampliamente usada en la literatura [31].

4.2. Procesamiento de las Imágenes

Se procederá a dividir el cien por ciento de las imágenes en tres: a) set de datos de entrenamiento (80%), b) set de datos de validación (10%) y, por último, el set de datos de prueba con el 10% restante. Para cada grupo se escogerán los datos cuidadosamente, dejando un ejemplo de cada uno de los tipos de lesiones que contenga el dataset.

El set de datos de prueba será usado solo para evaluar el modelo propuesto, esto implica que estos datos nunca serán utilizados para entrenar el modelo.

4.2.1. Aumento del Set de Datos de Entrenamiento

Al set de entrenamiento y validación se le aplicaran distintas técnicas de *data augmentation* como desplazamientos, rotaciones y giros aleatorios. Todas estas transformaciones se harán a pequeña escala, para no alterar la naturaleza de las imágenes. También se hará un análisis para observar la distribución de la muestra y normalizar los datos para que cada tipo de lesión tenga el mismo número de imágenes.

Por el ultimo se deberá cambiar el tamaño de las imágenes a 224x224 que es tamaño por defecto que usan la mayoría de las arquitecturas más populares vistas.

Esto permitirá aumentar el tamaño del set de datos de entrenamiento y validación.

4.3. Modelo

Se entrenará una arquitectura de DCNN para la clasificación lesiones en las ecografías mamarias y se analizara la precisión del modelo con distintas métricas de evaluación [32]. Esto se describe a continuación.

4.3.1. Arquitectura de DCNN y Entrenamiento

Se seleccionará y evaluará una arquitectura DCNN: en este caso de estudio se eligió *VGGNet-16*. Se seleccionó este modelo debido a que es una arquitectura poco profunda, en comparación con las más recientes, y no requiere gran capacidad de computo, ya que la disponibilidad de recursos es limitada para este proyecto.

Para llevar a cabo el entrenamiento se realizarán algunas modificaciones en la arquitectura original de VGGNet-16, pre entrenada sobre base de datos de Imagenet. Entre esas modificaciones están quitar el clasificador original. Luego se insertará un nuevo clasificador, acorde a la base de datos seleccionada. Por último, se procederá a entrenar el modelo.

5. Resultados Obtenidos

En la *tabla 3* que se muestra se muestran los resultados obtenidos. La precisión alcanzada por el modelo propuesto fue de un 62%.

Tabla 3 - Resultados

Tipo de Lesiones	Precision	Recall	f1-score	support
ARCH	0.00	0.00	0.00	5
ASYM	0.00	0.00	0.00	3
CALC	0.32	0.75	0.44	8
CIRC	0.20	0.33	0.25	6
MISC	0.14	0.25	0.18	4
NORM	0.86	0.48	0.62	52
SPIC	0.50	0.40	0.44	5
promedio / total	0.62	0.43	0.48	0.83

6. Conclusión y Trabajos Futuros

En nuestro estudio, se presentó un modelo de DL que implementa una arquitectura de DCNN para clasificar distintos tipos de lesiones directamente desde las ecografías mamarias. Demostró que es difícil alcanzar una alta precisión cuando no se posee una base de datos con gran número de ejemplos para cada tipo de lesión. Pero si se pudo observar cómo influyen las diferentes estrategias de entrenamiento para contrarrestar esta situación.

Futuros trabajos, podrían incluir explorar otras arquitecturas y probar este modelo con datasets que sean de mayor tamaño o utilizar múltiples fuentes de datos.

7. Formación de Recursos Humanos

El proyecto de investigación se encuentra estructurado de la siguiente manera:

- Director Consultor: Dra. Silvana Aciar
- Director Ejecutivo: Ing. Hugo Chanampe, se encuentra realizando su trabajo final de Maestría.
- Co-Director: Ing. Martin de la Vega, se encuentra realizando su trabajo final de Maestría.
- Integrantes:
  - Molinari Sotomayor, José Luis
  - Carrascosa, German
  - Lorefice, Alejandro Martin

Todos los integrantes son alumnos avanzados de las carreras de sistemas de la Universidad Nacional de La Rioja.

8. Referencias

[1] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[2] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, and D. Shen, "Deep convolutional neural networks for multi-modality isointense infant brain image segmentation," *Neuroimage*, vol. 108, pp. 214–224, 2015.

[3] "Insitituto Nacional del Cancer -Mortalidad," 2015. [Online]. Available: Mortalidad. [Accessed: 13-Apr-2018].

[4] R. a. Smith, V. Cokkinides, and H. J. Eyre, "American Cancer Society Guidelines for the Early Detection of Cancer, 2004," *CA. Cancer J. Clin.*, vol. 54, no. 1, pp. 41–52, 2004.

[5] Y. Bengio, A. Courville, P. V.-I. transactions on pattern, and undefined 2013, "Representation learning: A review and new perspectives," *ieeexplore.ieee.org*.

[6] S. Liu, S. Liu, W. Cai, S. Pujol, R. Kikinis, and D. Feng, "Early diagnosis of Alzheimer's disease with deep learning," in *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 2014, pp. 1015–1018.

[7] A. Birenbaum and H. Greenspan, "Longitudinal Multiple Sclerosis Lesion Segmentation Using Multi-view Convolutional Neural Networks," Springer, Cham, 2016, pp. 58–67.

[8] A. Akselrod-Ballin, L. Karlinsky, S. Alpert, S. Hasoul, R. Ben-Ari, and E. Barkan, "A Region Based Convolutional Network for Tumor Detection and Classification in Breast Mammography," Springer, Cham, 2016, pp. 197–205.

[9] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, no. December 2012, pp. 60–88, 2017.

[10] H. Cai, Y. Peng, C. Ou, M. Chen, and L. Li, "Diagnosis of Breast Masses from Dynamic Contrast-Enhanced and Diffusion-Weighted MR: A Machine Learning Approach," *PLoS One*, vol. 9, no. 1, p. e87387, Jan. 2014.

[11] M. Krishnan, S. Banerjee, … C. C.-E. S. with, and undefined 2010, "Statistical analysis of mammographic features and its classification using support vector machine," *Elsevier*.

[12] A. Krizhevsky, I. Sutskever, … G. H. information processing, and undefined 2012, "Imagenet classification with deep convolutional neural networks," *papers.nips.cc*.

[13] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, "Deep learning for visual understanding: A

review," *Neurocomputing*, vol. 187, pp. 27–48, 2016.

[14] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 142–158, 2016.

[15] J. S.-N. networks and undefined 2015, "Deep learning in neural networks: An overview," *Elsevier*.

[16] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998.

[17] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *Int. Conf. Learn. Represent.*, pp. 1–14, 2015.

[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015, vol. 07–12–June, pp. 1–9.

[19] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," Feb. 2015.

[20] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision." pp. 2818–2826, 2016.

[21] K. He, J. S.-P. of the I. conference on computer, and undefined 2015, "Convolutional neural networks at constrained time cost," *cv-foundation.org*.

[22] R. K. Srivastava, K. Greff, and J. Schmidhuber, "Highway Networks," May 2015.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[24] A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN features off-the-shelf: An astounding baseline for recognition," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 512–519.

[25] O. A. B. Penatti, K. Nogueira, and J. A. Dos Santos, "Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2015, vol. 2015–Octob, pp. 44–51.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, 2015.

[27] S. Ren, K. He, R. Girshick, J. Sun, J. Donahue, T. Darrell, J. Malik, W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, R. Girshick, D. Impiombato, S. Giarrusso, T. Mineo, O. Catalano, C. Gargano, G. La Rosa, F. Russo, G. Sottile, S. Billotta, G. Bonanno, S. Garozzo, A. Grillo, D. Marano, G. Romeo, D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 794, pp. 1–15, 2015.

[28] A. G. Howard, "Some Improvements on Deep Convolutional Neural Network Based Image Classification," *arXiv Prepr. arXiv1312.5402*, pp. 1–6, 2013.

[29] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *J. Mach. Learn. Res.*, vol. 15, pp. 1929–1958, 2014.

[30] "The mini-MIAS database of mammograms." [Online]. Available: http://peipa.essex.ac.uk/info/mias.html. [Accessed: 17-Aug-2018].

[31] N. Dhungel, G. Carneiro, and A. P. Bradley, "Automated Mass Detection in Mammograms Using Cascaded Deep Learning and Random Forests," *2015 Int. Conf. Digit. Image Comput. Tech. Appl.*, pp. 1–8, 2015.

[32] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, Jul. 2009.

Motor de búsqueda semántico de contenido académico en repositorios digitales bajo el protocolo OAI-PMH

José Federico Medrano[1], José Luis Alonso Berrocal[2], Carlos G. Figuerola[2]

jfmedrano@fi.unju.edu.ar, berrocal@usal.es, figue@usal.es

[1]VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

[2] REINA / Recuperación de Información Avanzada / Facultad de Traducción y Documentación

Universidad de Salamanca – España - C/ Francisco de Vitoria, 6-16

RESUMEN

A menudo las búsquedas de material académico que uno realiza no son del todo acertadas o no son tan exactas como uno quiere. Los resultados ofrecidos por los motores de búsqueda dependen en gran medida de los mecanismos internos utilizados y de los algoritmos de ordenación, tal es el caso de *Google Scholar* que emplea el *Page Rank* (Page, Brin, Motwani, & Winograd, 1999) para ordenar los resultados. Por otro lado, los repositorios digitales institucionales carecen de un buscador de material relacionado, puesto que los resultados de las búsquedas se basan en la existencia de algunos de los términos buscados en los campos de metadatos de los registros almacenados. Por esta razón, este proyecto propone la construcción de un metarepositorio que recolecte todos los registros de los repositorios digitales argentinos pertenecientes a instituciones educativas que implementan el protocolo *Open Archives Initiative Protocol for Metada Harvesting* y que permita realizar búsquedas semánticas de contenido relacionado a partir de una búsqueda inicial.

Palabras clave: *Buscador semántico; Repositorios digitales; OAI-PMH; PLN*

CONTEXTO

La línea de investigación aquí presentada se encuentra enmarcada dentro del Proyecto Consolidado D/B029 denominado "*Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino*", aprobado y financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy. Este proyecto es llevado a cabo en conjunto por dos grupos de investigación. En primer lugar liderado por el grupo de investigación VRAIn (Visualización y Recuperación Avanzada de Información) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, y en segundo lugar como colaboradores, el grupo REINA (Recuperación de Información Avanzada) de la Facultad de Traducción y Documentación de la Universidad de Salamanca

1. INTRODUCCIÓN

Son numerosas las instituciones y entidades que necesitan no solo preservar el material y las publicaciones que producen, sino también, estas

tienen como tarea publicar, divulgar y poner a disposición del público los resultados de la investigación y cualquier otro material científico-académico. Para este propósito existen los repositorios de libre acceso, que a través de iniciativas como la *Open Archives Initiative* (OAI) y de la aparición de instrumentos como el protocolo *Open Archives Initiative Protocol for Metada Harvesting* (OAI-PMH), facilitan esta tarea en gran medida.

El protocolo OAI-PMH propone unos lineamientos generales tanto para listar y recuperar (cosechar) metadatos de un repositorio (OAI Service Providers), como también para exponer recursos (OAI Data Providers) para que puedan ser cosechados por aplicaciones externas. Estos lineamientos proponen la organización de los recursos en conjuntos (sets), el uso del estándar XML para la representación y transporte de recursos (vía HTTP), y un conjunto de seis verbos necesarios para interactuar, como por ejemplo identificar el repositorio, listar conjuntos, listar formatos de metadatos soportados u obtener registros (Medrano, 2017)

La mayor parte del tiempo, los investigadores deben filtrar varios documentos académicos para encontrar aquellos relevantes para su investigación. Este filtrado muchas veces es engorroso y requiere emplear una considerable cantidad de tiempo. En la búsqueda de este tipo de material resulta útil contar con un listado de objetos relacionados no sólo con la temática buscada, sino también material que pueda estar relacionado semánticamente con el objeto de la búsqueda. Los repositorios institucionales (RI) poseen un motor de búsqueda interno, el cual ofrece un conjunto de resultados a partir de una búsqueda por palabra o frase, autor o fecha; buscando dentro de todas las colecciones o una colección en particular. Para los RI que utilizan como software de base a *DSpace*, el motor de búsqueda, dependiendo de la versión, es *Lucene* o SOLR o combinaciones de ambos. Estos motores permiten un gran número de opciones al momento de realizar la búsqueda (Prasad & Patel, 2005), pero son búsquedas sintácticas, es decir, se busca entre los registros almacenados la coincidencia de los términos ingresados realizando distintas combinaciones con los campos de metadatos de los mismos. Los resultados presentados, a partir de una búsqueda inicial, no ofrecen ninguna otra relación más allá de la existencia/coincidencia de algún término en común.

Ofrecer contenido semánticamente relacionado aportaría significativamente mejores resultados que búsquedas sintácticas, como fue comprobado en (Medrano, 2018). Por ello la necesidad de encarar este proyecto y extender el trabajo inicial mencionado aportando una herramienta capaz de ofrecer resultados "*más*" relevantes. La búsqueda semántica es una técnica de búsqueda de datos en la que una consulta de búsqueda tiene como objetivo no solo encontrar palabras clave, sino también determinar la intención y el significado contextual de las palabras que una persona está utilizando para la búsqueda.

La idea central es recolectar periódicamente los repositorios institucionales argentinos, para esto se consultará el ROAR (*Registry of Open Access Repositories*) que es una de las bases de datos de repositorios de libre acceso más grandes que existen, posee alrededor de 4734 repositorios registrados, con 57 repositorios pertenecientes a instituciones argentinas. El interés fue puesto en los repositorios que soportan el protocolo OAI-PMH, puesto que es un protocolo estándar para la recolección de metadatos y es el objeto de estudio de este trabajo.

La herramienta que se pretende desarrollar, una vez recuperados y almacenados los registros, se presentará como un sistema de recomendación

basado en contenido (Ricci, Rokach, & Shapira, 2011; Bai, y otros, 2019), en este tipo de sistemas el contenido desempeña un papel principal en el proceso de recomendación, en el que las calificaciones de los usuarios y las descripciones de los atributos de los elementos se aprovechan para hacer predicciones. La idea básica es que los intereses del usuario se puedan modelar sobre la base de las propiedades (o atributos) de los elementos que han calificado o accedido en el pasado. Los "elementos" suelen ser textuales, por ejemplo, correos electrónicos (Paik, y otros, 2001) o páginas web (Seroussi, 2010). La "interacción" generalmente se establece mediante acciones, como descargar, comprar, crear o etiquetar un artículo. Los elementos están representados por un modelo de contenido que contiene las características de los elementos. Las características suelen estar basadas en palabras, es decir, palabras sueltas, frases o n-*grams* (Beel, Gipp, Langer, & Breitinger, 2016).

Algunas de las aproximaciones relacionadas con sistemas de recomendación de contenido académico o relacionado a la educación están bien clasificadas y resumidas en (Beel, Gipp, Langer, & Breitinger, 2016), los autores mencionan algunos enfoques como sistemas de recomendación de libros, sistemas de recomendación educativa, servicios de alerta académica, búsqueda de expertos, resumen automático de artículos académicos, recomendadores de conjuntos de datos académicos y detección de plagio. En el mismo sentido, otras de las aproximaciones es el enfoque de (Son & Kim, 2018) el cual propone un sistema de recomendaciones para artículos académicos que combina el análisis de citas y el análisis de redes. Por otro lado, (Hwang, Wei, & Liao, 2010) plantea un sistema de recomendación basado en un esquema híbrido que emplea redes de coautoría y técnicas basadas en contenido.

El enfoque que aquí se propone consta de 2 etapas bien diferenciadas:

1. Proceso de recolección de información: el cual, como se mencionó previamente, consistirá en recolectar por completo, y de forma periódica, los RI Argentinos listados en ROAR y que soportan el protocolo OAI-PMH, es decir, aquellos que respetan el formato de metadatos Dublin Core[1] de manera obligatoria.
2. Diseño del modelo de predicción: para este proceso es necesario aplicar técnicas de Procesamiento del Lenguaje Natural (PLN) que permitan representar de forma adecuada la información y calcular/computar el grado de cercanía semántica entre una búsqueda inicial y el resto de registros almacenados.

Para el diseño del modelo de predicción, es necesario representar de algún modo la información, en este caso el conjunto de registros, para esto los campos utilizados en este estudio serán los títulos (campo *title*) y las descripciones (campo *description*, empleado para almacenar el resumen del elemento digital) por ser los más representativos de cada registro.

Para este trabajo se decidió utilizar la librería *Gensim* (Rehurek & Sojka, 2010) del lenguaje *Python*[2]. La elección del lenguaje y librería responden al hecho de que en el ámbito de ciencia de datos, *Python* es el lenguaje más versátil y quien mejor desempeño ofrece, además existe un gran número de librerías que reducen enormemente las tareas de desarrollo al incorporar implementaciones de los algoritmos y técnicas más utilizados. En este sentido las técnicas de representación de documentos seleccionadas para el desarrollo de los

[1] http://dublincore.org/

[2] https://www.python.org/

experimentos fueron: *Term Frequency Inverse Document Frequency* (TF-IDF) (Aizawa, 2003), *Latent Semantic Indexing/Latent Semantic Analysis* (LSI/LSA) (Landauer & Dumais, 1997; Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) y *Word Mover's Distance* (WMD) (Kusner, Sun, Kolkin, & Weinberger, 2015; Mikolov, Sutskever, Chen, Corrado, & Dean, 2013). Para cada una de las técnicas mencionadas se diseñará un modelo de predicción el cual, tomando el título y la descripción de un resultado elegido, ofrecerá un conjunto de publicaciones relacionadas semánticamente con el mismo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de trabajo principal de este proyecto de investigación es el estudio, evaluación e implementación de técnicas de representación de documentos (PLN) que trabajarán sobre los registros recolectados de RIs para ofrecer un motor de búsqueda semántico de material académico. Este acercamiento pretende mejorar los procesos de búsqueda y filtrado que son llevados a cabo al momento de realizar una búsqueda bibliográfica.

Abarca las siguientes áreas y disciplinas: Recuperación de Información, Bases de Datos, Desarrollo Web, Inteligencia Artificial, PLN, entre otras.

3. RESULTADOS OBTENIDOS/ESPERADOS

En una primera instancia se espera emplear las técnicas de representación de documentos elegidas y evaluar los resultados entregados tanto de forma manual como comparándolos con motores de búsqueda actuales. Para luego seleccionar la técnica o combinación de estas que mejor resultados entregue.

Una vez que los resultados demuestren ser fiables y se ajusten a ciertos parámetros de calidad, se espera poder publicar la herramienta para que sea accesible y utilizada de forma libre por la comunidad académica.

4. FORMACIÓN DE RECURSOS HUMANOS

Los autores de este trabajo han desarrollado investigaciones sobre la temática presentada. El director del mismo, el Dr. J. Federico Medrano, se encuentra realizando investigaciones postdoctorales relacionadas con bases de datos bibliográficas de libre acceso, entre las que se encuentran los Repositorios Institucionales como una fuente adicional de material científico-académico. Actualmente es becario postdoctoral de la Fundación Carolina[3] y realiza su investigación en conjunto con la Universidad de Salamanca, donde presentan su apoyo y colaboración el Dr. José Luis Alonso Berrocal y el Dr. Carlos G. Figuerola.

Por otro lado, en este proyecto formarán parte alumnos del último año de la carrera Ingeniería Informática de la Universidad Nacional de Jujuy en donde imparte docencia el director. De los cuales se espera puedan desarrollar su trabajo fin de carrera con temáticas afines.

5. BIBLIOGRAFÍA

Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing & Management, 39*(1), 45-65.

[3] https://www.fundacioncarolina.es/

Bai, X., Wang, M., Lee, I., Yang, Z., Kong, X., & Xia, F. (2019). Scientific Paper Recommendation: A Survey. *IEEE Access, 7*, 9324--9339.

Beel, J., Gipp, B., Langer, S., & Breitinger, C. (2016). Research-paper recommender systems: a literature survey. *International Journal on Digital Libraries, 17*(4), 305-338.

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science, 41*(6), 391.

Hwang, S., Wei, C., & Liao, Y. (2010). Coauthorship networks and academic literature recommendation. *Electronic Commerce Research and Applications, 9*(4), 323-334.

Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). From word embeddings to document distances. *International Conference on Machine Learning*, (pp. 957-966).

Landauer, T., & Dumais, S. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review, 104*(2), 211.

Medrano, J. F. (2017). Calidad en repositorios digitales en argentina, estudio comparativo y cualitativo. *VII Conferencia Internacional BIREDIAL-ISTEC'17 y XII SIBD.* La Plata.

Medrano, J. F. (2018). Filtrado basado en contenido para artıculos académicos en repositorios institucionales. *Congreso Argentino de Ciencias de la Computación (CACIC).* Tandil.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Neural Information Processing Systems*, 3111-3119.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The pagerank citation ranking: Bringing order to the web.* Stanford InfoLab.

Paik, W., Yilmazel, S., Brown, E., Poulin, M., Dubon, S., & Amice, C. (2001). Applying Natural Language Processing (NLP) based metadata extraction to automatically acquire user preferences. *Proceedings of the 1st international conference on Knowledge capture.*, (pp. 116-122).

Prasad, A., & Patel, D. (2005). Lucene search engine: An overview. *DRTC-HP International.*

Rehurek, R., & Sojka, P. (2010). Software Framework for Topic Modelling with Large Corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks.* (pp. 45-50). Valletta, Malta: ELRA.

Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. *Recommender systems handbook*, 1-35.

Seroussi, Y. (2010). Utilising user texts to improve recommendations. *International Conference on User Modeling, Adaptation, and Personalization* (pp. 403-406). Springer.

Son, J., & Kim, S. (2018). Academic paper recommender system using multilevel

simultaneous citation networks.
Decision Support Systems, 105, 24-33.

RNA-AP: Redes Neurales Artificiales con Aprendizaje Profundo

Juan C. Vázquez[1], Leticia Constable[1]

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

[1] {jcjvazquez, leticiaconstable}@gmail.com

RESUMEN

Este proyecto se adentrará en las técnicas de aprendizaje profundo (*Deep Learning*) de redes neurales artificiales, que en los últimos años ha producido importantes avances en la Inteligencia Artificial a nivel mundial. El aprendizaje profundo ha tenido alto y positivo impacto en aplicaciones de reconocimiento de imágenes, pero también ha tenido resonantes fracasos debido a que no se entiende con claridad su funcionamiento interno. El proyecto intentará, por un lado obtener el conocimiento necesario para desarrollar software de experimentación y aplicación, y por otro, entender la teoría subyacente, el por qué esta técnica funciona mejor que las anteriores, cuáles son sus bondades y limitaciones, en qué campos es aplicable y en cuáles no, y en lo posible, hacer algún aporte significativo a la teoría.

Palabras clave: redes neurales artificiales, aprendizaje profundo, aplicaciones.

CONTEXTO

El proyecto es parte del recientemente conformado Grupo GA^{2}LA (Grupo de Investigación, Desarrollo y Transferencia en Aprendizaje Automático, Lenguajes y Autómatas) de la Facultad Córdoba de la UTN.

Desde el año 2004, en el Laboratorio de Investigación de Software del Departamento de Ingeniería en Sistemas de Información, sede de GA^{2}LA, se ha desarrollado una secuencia de proyectos sobre redes neurales artificiales que han generado productos y aplicaciones que han sido transferidos (muchos de ellos en cooperación con investigadores de CONICET) a numerosos centros de investigación en diversas locaciones en América Latina, a empresas comerciales y a gobiernos municipales de Argentina y Brasil. En estos proyectos se han utilizado redes neurales artificiales de perceptrones multicapa con entrenamiento por retropropagación de errores, con aplicaciones a las ciencias sociales, a los negocios y a la ingeniería de software.

1. INTRODUCCIÓN

La Inteligencia Artificial como rama de estudio dentro de la Informática, ha buscado desde sus orígenes lograr un comportamiento inteligente de las computadoras mediante gran cantidad de métodos y técnicas que pueden englobarse básicamente en dos estrategias principales:

a) El llamado ***enfoque simbólico***, intenta emular la forma en la cual se cree que las personas piensan, utilizando diversas versiones de la Lógica (proposicional, de predicados, rebatible, modal, temporal, difusa, etc.) para especificar hechos y relaciones causales entre ellos, que representen la realidad en un dominio determinado. Luego, a través procesos de inferencia con rígida justificación formal, que usen razonamiento deductivo (monótono o no monótono) y/o inductivo, se puede interrogar al modelo del mundo así generado, para obtener *respuestas inteligentes* a los problemas planteados. Básicamente este abordaje supone que la *inteligencia* surge de extensas bases de conocimiento y razonamientos que las utilizan. Los sistemas expertos son productos típicos generados por este enfoque.

b) La segunda estrategia suele denominarse ***enfoque sub-simbólico***; aquí se engloban gran cantidad de teorías, modelos, técnicas y herramientas que emulan la forma en que la naturaleza ha resuelto a lo largo de milenios los problemas naturales, ya sea si-

mulando su método de *selección natural* (computación evolucionaria, autómatas celulares, programación evolutiva, algoritmos genéticos y otros) o modelando matemática y computacionalmente sus soluciones más relevantes (sistemas conexionistas, sistemas inmunológicos, estrategias de enjambre, etc.).

Una de las ventajas del enfoque simbólico es que puede explicar cómo arriba a sus conclusiones, lo que en algunos dominios como la medicina es de fundamental importancia para que los usuarios adopten como buenas las soluciones logradas y confíen en ellas.

Dentro del enfoque sub-simbólico, destacan los sistemas conexionistas que emulan la arquitectura del sistema nervioso de los animales, y en particular, de su órgano principal: el cerebro; una inmensa cantidad de unidades simples de procesamiento de señales (las neuronas) intrincadamente conectadas enviándose mensajes y colaborando para soportar el pensamiento. Cabe aclarar aquí, que los modelos matemáticos de las redes neuronales que componen el cerebro y sus partes, son una metáfora muy simplificada del complicadísimo mecanismo biológico real, el cual aún no termina de comprenderse por completo. Se piensa bajo este esquema, que la *inteligencia* es una *característica emergente* de esta complejidad extrema puesta a funcionar.

Los modelos de redes neurales artificiales (RNA) y las técnicas asociadas, se engloban actualmente en lo que ha dado en llamarse ***aprendizaje automático*** (*machine-learning* en inglés) e intervienen en muchos dispositivos y procesos usuales de la sociedad moderna (búsquedas de Internet, teléfonos inteligentes, traducción automática de lenguaje escrito y hablado, filtrado de contenidos en redes sociales, reconocimiento de imágenes y su clasificación, por nombrar solo algunos).

El interés por las RNA en la Inteligencia Artificial, ha pasado históricamente por tres etapas distintas:

1) La primera ola inicia en a fines de la década de 1950 con la presentación del "perceptron" por Rosenblat, un modelo de red de dos capas de neuronas del tipo McCulloch-Pitts, que se desempeña como clasificadora de patrones simples.
2) La segunda comienza en 1982 con las redes autoasociativas de Hopfield, utilizando aprendizaje hebbiano, y por la propuesta de un algoritmo general de aprendizaje (Rumelhart-1986) sobre redes multicapa de perceptrones con propagación unidireccional: la retropropagación de errores (*back-propagation* en inglés); intensa actividad de investigación y desarrollo siguen a estas propuestas, generándose gran cantidad de modelos distintos de neuronas, arquitecturas de conexionado y métodos de funcionamiento, con aplicaciones más o menos destacadas.
3) La tercera ola, en la cual la comunidad IA considera que actualmente nos encontramos, surge de estudios realizados por Hinton y su equipo publicados en 2006 en la revista *Science*, sobre una jerarquía de representaciones de datos (desde las de más bajo nivel morfológico, hasta abstracciones de mayores niveles) logradas a través de *autocodificadores,* que aportan luz sobre cómo están haciendo su trabajo las RNA durante un proceso de reconocimiento de imágenes, y que mejoran el comportamiento del algoritmo de retropropagación de errores enormemente. Estos hallazgos dan nacimiento al ***aprendizaje profundo*** **(***Deep Learning* en inglés), que se consolida como tecnología emergente en el campo de las redes neurales artificiales con una serie de trabajos del mismo Hinton (hoy en Google) y de LeCun (hoy en Facebook) que les valen numerosos premios en competiciones de IA a nivel mundial, desde finales de 2010.

A partir de aquí los avances en los últimos pocos años son tan rápidos, y los éxitos obtenidos en igualar capacidades reservadas hasta ahora solo a los humanos son tan impresionantes, que empresas como Google, Microsoft, Facebook y otras han iniciado programas de investigación a gran escala y desarrollo de aplicaciones, bajo este nuevo paradigma de aprendizaje profundo. Incluso hace no mucho, Google ha hecho de dominio

público una librería de software (*TensorFlow*), para implementar redes neurales con aprendizaje profundo.

En la Facultad Córdoba de la UTN, iniciamos en 2004 un proyecto sobre el estudio de relaciones que suponíamos debían existir entre el funcionamiento del algoritmo de aprendizaje por retropropagación de errores en redes neurales multicapa de perceptrones, y la evolución espacio-temporal de los autómatas celulares unidimensionales; también se ensayaron redes de Hopfield siguiendo la evolución de los patrones de activación de sus elementos durante su reconocimiento de caracteres. El proyecto 25/E078 tuvo moderado éxito, como resulta del informe técnico presentado en 2005, ya que se encontró la relación buscada sólo durante el entrenamiento de un tipo particular de problemas de conversión de números entre bases numéricas, bajo una codificación de las seis ideadas, de la evolución sufrida por los pesos sinápticos durante el aprendizaje de la red.

El conocimiento logrado durante este proyecto y el software desarrollado, propició la generación de una línea de investigación que se consolidó con el desarrollo de nuevos proyectos siguiendo los temas planteados (proyectos 25/E128 y 25/E174) y sobre nuevas temáticas y aplicaciones de las RNA y de técnicas asociadas (prevención de incendios forestales en la provincia de Córdoba, cálculo del índice de riesgo para la salud de la vivienda urbana, determinación de implicaciones textuales dentro de documentos escritos, y otros).

Asimismo se establecieron convenios de cooperación con el Instituto de Investigaciones Geo-Históricas de CONICET, con grupos de otras universidades y con centros de I+D; se desarrollaron entrenadores de redes neuronales y redes ya entrenadas para su aplicación en ciencias sociales y en la industria local de software, desarrollos que fueron documentados, se obtuvo sobre ellos derechos de autor y se transfirieron a empresas, centros de investigación en Argentina, Brasil, Paraguay, Cuba, Colombia y a gobiernos municipales de Argentina y Brasil para su aplicación. También los proyectos sirvieron de base para la concreción de tesis de grado de becarios alumnos, para la introducción a la investigación de becarios graduados, y se generaron ideas sobre tesis de especialidad, maestría y doctorado, algunas de las cuales ya finalizaron exitosamente y otras que aún están en desarrollo.

Todo esto se concretó *"montados en la segunda ola"* de interés de la Inteligencia Artificial sobre las redes neurales artificiales.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El grupo GA^{2}LAtiene como ejes para su funcionamiento líneas de investigación, a saber:

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

El proyecto presentado aquí en enmarca en la tercera línea de I+D.

3. RESULTADOS OBTENIDOS / ESPERADOS

A la luz de los avances y éxitos de las nuevas técnicas de aprendizaje profundo ya comentadas, en el presente proyecto queremos indagar sobre los conceptos y herramientas que están dando lugar a la tercera ola de interés sobre las RNA, lograr conocimiento teórico-práctico sobre el tema y desarrollar software que lo implemente, con distintos objetivos a saber:

a) Académicos: la temática es interesante por sí misma y está relacionada directamente con las cátedras de Inteligencia Artificial de la carrera de Ingeniería en Sistemas de Información y con la de Análisis Numérico y Cálculo Avanzado de la carrera de Ingeniería Industrial; además se relaciona indirectamente con temáticas de las cátedras de Sintaxis y Semántica del Lenguaje, de Matemática Discreta y de Análisis Matemático. Los integrantes del equipo de investigación propuesto participamos en algunas de estas cátedras.
b) Investigación: se pretende actualizar el conocimiento que ya tenemos sobre redes

neurales y entender los fundamentos de las técnicas agrupadas en el aprendizaje profundo, y la forma de aplicarlos.

c) Transferencias: en anteriores proyectos se han desarrollado productos de redes neurales artificiales tradicionales, que han sido transferidos; los usuarios están actualmente solicitando actualizaciones, modificaciones y capacitación sobre estos productos para producir mejoras, internacionalizarlos y adaptarlos a nuevas aplicaciones reales. Por otro lado, otros proyectos actualmente en desarrollo en nuestra unidad académica que usan redes neurales tradicionales (dirigidos por anteriores becarios e integrantes de nuestros anteriores proyectos), colaborarán con el presente y serán también receptores de estos nuevos conocimientos y herramientas desarrolladas.

Con estos objetivos en mente, el proyecto intentará:

- Obtener conocimiento sobre las técnicas de aprendizaje profundo: sobre sus fundamentos, sobre qué tipos de arquitecturas de redes neurales son aplicables, qué modelos de neuronas y funciones de activación utilizan, para qué tipo de tareas son útiles y para cuales no lo son, y cómo funcionan los algoritmos de aprendizaje por retropropagación de errores en cascadas de autocodificadores u otros modelos sugeridos para su implementación.
- Identificar y comprender el significado de la *jerarquía de representaciones de datos* que obtiene el aprendizaje profundo, cómo funciona el proceso de extracción de características no supervisada de los datos (desde las de más bajo nivel en las primeras capas de las redes, hasta las abstracciones de más alto nivel en las superiores) que sustentan esa jerarquía y la forma de visualizarlas.
- Desarrollar software que implemente el aprendizaje profundo para problemas diseñados de testeo y establecer cómo determinar su efectividad y la de sus posibles alternativas, para comparación.
- Generar software de entrenamiento de redes neurales artificiales con aprendizaje profundo de propósito general y especializarlo, de ser posible, para el caso particular de la metodología DRVS del IIGHI-UNNE-CONICET y desarrollar un módulo de producción RVS con las redes ya entrenadas, para actualizar las transferencias realizadas.

4. FORMACIÓN DE RECURSOS HUMANOS

Se prevé tener durante el transcurso del proyecto becarios alumnos de grado, prácticas supervisadas de alumnos de la carrera de Ingeniería en Sistemas de Información y eventualmente becarios graduados. Sobre los becarios alumnos, se hará un seguimiento de su desempeño académico tratando de apoyarlos en su carrera y asegurando que el proyecto no influya negativamente sobre la misma.

El proyecto será ofrecido a la Maestría en Ingeniería en Sistemas de Información de la Facultad Córdoba de UTN y al Doctorado en Ingeniería en Sistemas de Información, para servir de asiento de posibles tesis. El profesor Juan Giró, integrante consultor del proyecto, es docente de la Maestría citada.

Se prevé para todos los integrantes, efectuar capacitación interna respecto de los temas bajo investigación y se procurarán becas para los mismos, para que anualmente puedan acceder al menos a un curso de posgrado.

Es intención del proyecto, generar ideas para una nueva tesis de doctorado en Ingeniería en Sistemas de Información que tiene intenciones de desarrollar la Ing. Constable. Por otro lado, se prevé que la temática de investigación impulse la culminación de la tesis doctoral en curso del Ing. Arcidiácono sobre reconocimiento de imágenes (*Reconocimiento e Identificación de Patrones Morfológicos de Fibra Textil mediante un algoritmo de Memoria Jerárquica Temporal*) y de la tesis de maestría, también en curso, del Esp. Ing. Gualpa sobre ingeniería de software (*La planificación de entregas (release planning) en proyectos de desarrollo y mantenimiento de software, soportada por*

métodos cuantitativos multicriterio para la toma de decisiones).

5. BIBLIOGRAFÍA

1. Anderson J. (2007), *Redes Neurales*, Alfaomega Grupo Editor, D.F., México.
2. Araujo B. (2006), *Aprendizaje Automático: conceptos básicos y avanzados*, Pearson, Madrid, España.
3. Del Brío M., Sanz Molina A. (2006), *Redes Neuronales y Sistemas Borrosos*, Ra-Ma., Zaragoza, España.
4. Deolalikar V. (2002), *A Two-Layer Paradigm Capable of Forming Arbitrary Decision Regions in Input Space*, IEEE Transactions on Neural Network, Vol. 13, #1, pp. 15-21.
5. Hinton G., Salakhutdinov R. (2006), Reducing the Dimensionality of Data with Neural Network, Science, Vol. 313, pp. 504-507.}
6. Krizhevsky A., Sutskever I., Hinton G. (2017), *ImageNet Classification with Deep Convolutional Neural Network*, Communications of ACM, June 2017, Vol. 60 – Nr 6 – pp. 84-90, U.S.A.
7. LeCun Y., Bengio Y., Hinton G. (2015), *Deep Learning*, Nature, Vol. 521, pp. 436-444.
8. LeCun Y., Botton L., Genevieve B., Müller K. (1998), *Efficient BackProp*, en Neural Network: tricks of the trade, Springer, Berlin, Germany.
9. Rojas, Meichtry, Ciuffolini, Vazquez, Castillo (2008), *Repensando de manera holÍstica el riesgo de la vivienda urbana precaria para la salud: un análisis del enfoque de la vulnerabilidad sociodemográfica*, Revista Salud Colectiva, Vol. 4 / Nro. 2, Buenos Aires, Universidad Nacional de Lanús y Asociación Civil Salud Colectiva Centro de Estudios para la Salud.
10. Rumelhart D. y McClelland J. (1986), *Parallel Distributed Processing*, Vol. 1: *Foundations*, MIT Press.
11. Tang J., Deng C. y Huang G., *Extreme Machine Learning for Multilayer Perceptron*, IEEE Transaction on Neural Network and Learning Systems, Vol. 27, Nr. 4, pp. 809-821, April-2016, IEEE.
12. Vázquez, J.C. (2009), *Unificación de Modelos Evolutivos*, Congreso de Inteligencia Computacional Aplicada, Universidad de Palermo, Buenos Aires, Argentina.
13. 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.
14. 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.

Una ruta inteligente y sostenible
Inteligencia aplicada a la recolección de residuos tecnológicos

D. Pandolfi[1], J. Rasjido[1], A. Villagra[1], S. Orozco[1], D. Perez[1], N. Serón[1], Varas, G. Leguizamon[12]
{jrasjido,dpandolfi,avillagra,sorozco,dperez,nseron,vvaras}@uaco.unpa.edu.ar,
legui@unsl.edu.ar

[1] Laboratorio de Tecnologías Emergentes (LabTEm)
Instituto de Tecnología Aplicada (ITA) - 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 gestión de los desechos Tecnológicos, es un problema emergente, así como una oportunidad económica, dados los volúmenes de residuos que se generan y el contenido de materiales valiosos que lo componen. La recolección, transporte, tratamiento, reciclaje y eliminación de residuos tecnológicos requiere de soluciones económica y ambientalmente sostenibles. Las soluciones a dichos retos, demandan de enfoques inter-displinarios que incluyen aspectos ambientales, tecnológicos económicos, sociales y legislativos. El paradigma de la economía circular plantea que los residuos al final del ciclo de vida útil de un producto pueden ser materias primas para unos nuevos productos.

La programación de la logística de recolección de residuos tecnológicos corresponde al problema de enrutamiento de vehículos (*Vehicle Routing Problem*, VRP) que presentan una extensa disponibilidad bibliográfica de los problemas de enrutamiento verde de vehículos (GVRP). Los problemas de enrutamiento de VRP en logística describen la complejidad en los casos del mundo real.

Particularmente, el propósito de esta línea trabajo es mostrar como la aplicación de un algoritmo basado en inteligencia colectiva mejora la planificación de recolección de residuos tecnológicos para un caso de estudio local sobre con diferentes escenarios e instancias de recolección contenedores de residuos y vehículos con diferentes capacidades.

Palabras clave: Ciudades Inteligentes, Residuos Tecnológicos, Problema de enrutamiento de vehículos, Inteligencia Computacional, ACO.

1. 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/B225 "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).

2. INTRODUCCIÓN

Una ciudad inteligente es una ciudad donde el desempeño de sus componentes fundamentales (economía, movilidad, medioambiente, salud, gobernanza, etc.), están construidas sobre la base de nuevas capacidades donde sus ciudadanos pueden afrontar de manera inte-

ligente los desafíos futuros de sostenibilidad urbana. El crecimiento demográfico urbano constante de los próximos años, hace prever que la sociedad y los gobiernos locales se enfrentarán a nuevos retos ante el constante incremento de la generación de residuos tecnológicos.

La creciente demanda en el uso las tecnologías de la información y comunicación (TIC) han revolucionado nuestra vida cotidiana, sin embargo, han contribuido a aumentar las preocupaciones ambientales y sociales. La aceleración de factores como el agotamiento y escasez de recursos, daños ambientales, uso de agua y energía y un incremento insostenible de desechos tecnológicos principalmente en zonas urbanas, se ha convertido en una preocupación para la industria y el gobierno. El sostenimiento del ritmo de consumo a mediano y largo plazo depende de estrategias que permitan el reciclaje de los desechos y el fortalecimiento de una economía circular que realice un cambio sistémico, en la investigación y la innovación, tanto tecnológica como logística. La recuperación de residuos tecnológicos requiere de estrategias inteligentes de recolección que minimicen los costos asociados, maximicen el reuso de materias primas y la reparabilidad de los bienes para la prolongación de su vida útil.

Los equipos eléctricos y electrónicos (EEE) se convierten en uno de los grupos más importantes de residuos que contienen materiales fáciles de reciclar. Los desechos electrónicos son conocidos como RAEE (en español, Residuos de Aparatos Eléctricos y Electrónicos), o *e-waste*, o bien WEEE (en inglés, *Waste Electrical and Electronic Equipment*). Sin embargo, contienen muchas sustancias que son tóxicas y potencialmente peligrosas para el medio ambiente y la salud humana (Oguchi et al., 2012). La reducción, reciclaje y eliminación de los desechos finales de los WEEE se ha convertido en uno de los principales objetivos para el desarrollo sostenible de la creciente industria tecnológica. La tasa de WEEE está creciendo a un ritmo alarmante, especialmente en los países desarrollados, donde los mercados están saturados con grandes cantidades de nuevos productos electrónicos (Widmer et al., 2005). En 2013, un informe de StEPI (*Solving the e-waste Problem Initiative*), una iniciativa financiada por Naciones Unidas alertó que la cifra de WEEE se había elevado hasta cerca de los 49 millones de toneladas, un promedio de 7 kilos por cada uno de los 7.000 millones de habitantes del planeta Tierra. A la luz de estos datos, no resulta muy difícil hacerse una idea de la importancia que adquiere el reciclaje de este tipo de residuos.

Las empresas de recolección de residuos aplican diversos métodos de recolección, y pueden dividirse en estacionarios y móviles (Baxter et al., 2016). El objetivo tradicional de la gestión de distribución o recolección de productos consiste de minimizar los costos de todo el sistema. *Green Logistics* (GL) se ha convertido en una tendencia en la gestión de la distribución de bienes y la recolección de productos al final de su vida útil. Con su enfoque en maximizar el valor económico y ambiental mediante el reciclaje y el control de emisiones, GL contribuye al desarrollo sostenible de la industria, pero también requiere un esquema de transporte más completo cuando se realizan servicios de logística (Lin et al. 2014 (a) (b)).

Boussaïd et al. 2013, desarrollan un estudio de varios algoritmos de optimización inspirados en las metáforas del comportamiento de enjambre en la naturaleza, conocida como Inteligencia Colectiva, (SI, *Swarm Intelligent*). Ejemplo de SI son Optimización de Colonias de Hormigas (ACO, *Ant Colony Optimization*), la Optimización de Cúmulo de Partículas (PSO, *Particle Swarm Optimization*), la Optimización de Forrajeo Bacteriano (BFOA, *Bacterial Foraging Optimization Algorithm*), Optimización de Colonias de Abejas (BCO, *Bee Colony Optimization*), Sistemas Inmunes Artificiales (AIS, *Artificial Immune Systems*) y Optimización Basada en Biogeografía (BBO, *Biogeography-Based Optimization*). En particular, ACO fue propuesta fue propuesta por Dorigo et al (1996, 1991)

como una meta-heurística inspirada en el comportamiento de una colonia de hormiga en el proceso de forrajeo y aplicada principalmente para la solución de problemas de optimización combinatoria.

El problema de recolección estacionaria se lleva a cabo en puntos de recolección específicos de desechos, ubicados en lugares populares o en los alrededores de los lugares de residencia. La elección de la ubicación de dichos centros y la elección de un recorrido óptimo, representa distintos problemas de decisión. Así mismo, la recolección móvil, propone nuevos medios de recolección de WEEE a pedido y principalmente está asociado a residuos tecnológicos medianos o grandes (Nowakowski et al., 2017).

3. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de trabajo se enfoca en la generación de rutas recolección que permitan minimizar ciertos factores económicos y ambientales y maximizar los beneficios referidos a al reciclaje y tratamiento de WEEE. De los objetivos económicos podemos mencionar: minimizar el tiempo de recolección, kilómetros recorridos, maximizar el ahorro de combustible en los vehículos, minimizar la cantidad de vehículos, todo lo cual llevaría a obtener menores costos. Un objetivo económico adicional está dado por la valoración de la recuperación las materias primas tales como material ferroso, cobre, y metales preciosos, etc. De los beneficios ambientales podemos mencionar: la disminución en la emisión de gases efecto invernadero (CO_2, CH_4 y N_2O) y los gases de contaminación ambiental (CO, SO_2, PM, NOx). Tanto los objetivos económicos como ambientales promueven una mejor calidad de servicio e imagen, para las empresas y los gobiernos locales.

4. 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 Inteligencia Computacional para resolver problemas de VRP. (b) Proponer instancias de problemas relacionados a recolección de WEEE. (c) Proponer y seleccionar elementos de medición relacionados a los problemas identificados. (d) Diseñar y construir prototipos de software relacionados a los problemas.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo se encuentra formado por cuatro investigadores con distintos niveles de posgrado, dos Doctores en Ciencias de la Computación, dos Magisters en Ciencias de la Computación, tres Ingenieros en Sistemas y un estudiante de la Carrera Ingeniería en Sistemas.

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.

BIBLIOGRAFÍA

J. BAXTER, K.A., LYNG, C. ASKHAM, O.J. HANSSEN (2016). “High-quality collection and disposal of WEEE: environmental impacts and resultant” issues. Waste Manage. 57, 17–26.

I BOUSSAÏD, J. LEPAGNOT, P. SIARRY, A survey on optimization metaheuristics, Information Sciences, Volume 237, 2013, Pages 82-117, ISSN 0020-0255, https://doi.org/10.1016/j.ins.2013.02.041.

M. DORIGO, V. MANIEZZO, A. COLORNI, (1991) Positive Feedback as a Search Strategy, Technical Report 91-016, Dipartimento di Elettronica, Politecnico di Milano, Milan, Italy, 1991.

M. DORIGO, V. MANIEZZO, A. COLORNI, (1996) The ant system: optimization by a colony of cooperating agents, IEEE Transactions on Systems, Man, and Cybernetics Part B26 29–41.

C. LIN, K.L. CHOY, G.T.S. HO, T.W. Ng (a), A Genetic Algorithm-based optimization model for supporting green transportation operations, Expert Systems with Applications, Volume 41, Issue 7, 2014, Pages 3284-3296, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2013.11.032.

C. LIN, K.L. CHOY, G.T.S. HO, S.H. CHUNG, H.Y. LAM (b), Survey of Green Vehicle Routing Problem: Past and future trends, Expert Systems with Applications, Volume 41, Issue 4, Part 1, 2014, Pages 1118-1138, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2013.07.107.

P. NOWAKOWSKI, A. KRÓL, B. MRÓWCZNSKA (2017), Supporting mobile WEEE collection on demand: A method for multi-criteria vehicle routing, loading and cost optimisation, Waste Management, Volume 69, 2017, Pages 377-392, ISSN 0956-053X, https://doi.org/10.1016/j.waman.2017.07.045.

M. OGUCHI, H. SAKANAKURA, A. TERAZONO (2012). Toxic metals in WEEE: characterization and substance flow analysis in waste treatment processes. Sci. Total Environ. http://dx.doi.org/10.1016/j.scitotenv.2012.07.078.

R. WIDMER, H. OSWALD-KRAPF, D. SINHA-KHETRIWAL, M., SCHNELLMANN, H., BÖNI (2005). Global perspectives on e-waste. Environmental Impact Assessment Review 25 (5), 436–458.

Arquitecturas, Redes

y Sistemas Operativos

Adaptación de Driver Serial para la placa STM32F103C8T6 y su utilización en controladores de GSM – Bluetooth

Waldo Valiente, Esteban Carnuccio, Mariano Volker, Graciela De Luca, Gerardo García, Raúl Villca, Germán Lorenz, Matias Adagio
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Dirección: Florencio Varela 1703 – CP 1754 –
{wvaliente, ecarnuccio, mvolker, gdeluca, ggarcia} @unlam.edu.ar,
{raul.villcasd, germangelv, mati.adagio} @gmail.com

RESUMEN

En la actualidad la gran mayoría de los proyectos orientados a Internet de las Cosas, utilizan sistemas embebidos en donde resulta necesario emplear comunicación serial para comunicar la plataforma seleccionada con los distintos componentes de hardware que conforman el dispositivo IoT. Existen diversas cantidades de sensores, actuadores y mecanismos de comunicación que emplean el protocolo serial para su correcto funcionamiento. Las bases del sistema embebido que se está realizando en esta investigación se sustentan en este principio. En consecuencia, el presente trabajo expone de qué manera se desarrolló un módulo serial que permitiese utilizar la plataforma STM32F103C8T6 con dispositivos de este tipo. A partir de la utilización de este módulo, se explica de qué forma se generó un Controlador GSM, que permite manejar el SIM800L a través de comandos AT, de forma no bloqueante en un único hilo de ejecución. De la misma manera se detalla cómo se desarrolló un Controlador Bluetooth, que otorga la posibilidad de monitorear y utilizar en forma remota el dispositivo IoT durante la etapa de desarrollo. Por ese motivo en el presente trabajo se explican los métodos empleados en dicha construcción, detallando el algoritmo utilizado.

Palabras clave: *STM32, bibliotecas HAL, Driver Serial, SIM800L, HM-10.*

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 CYTMA2 de la Universidad Nacional de La Matanza, el cual está 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 y en Internet de las cosas.

1. INTRODUCCIÓN

Esta investigación pretende desarrollar un dispositivo que ayude a realizar esta tarea de asistencia remota empleando el concepto de Internet de las Cosas, de acuerdo a lo expuesto en [1]. Implementando un dispositivo IoT que permita la asistencia de personas mediante monitoreo remoto. Debido a que el sistema embebido IoT, debe ser fácilmente portable para

que el usuario lo pueda utilizar en su vida cotidiana, el tamaño del dispositivo y su consumo energético son factores preponderantes. En consecuencia a partir de un extenso estudio para poder cumplimentar estos requisitos de hardware adecuadamente, se ha seleccionado la arquitectura Cortex-M3, a través de la utilización de la placa STM32F10C8T6, también conocida como Blue Pill. [2]. El dispositivo IoT que se está desarrollando utilizará distintos sensores, actuadores y mecanismos de comunicación. Entre los principales sensores y actuadores que se están utilizando se encuentran: el acelerómetro MPU6050 y el GPS C3-470. No obstante una de las partes elementales del dispositivo IoT es la parte comunicacional con componentes externos; tales como un servidor en la nube, computadoras y otros dispositivos Bluetooth. Para ello se está utilizando el SIM800L V2, para la comunicación GSM/GPRS con un servidor, además del Bluetooth 4.0 HM-10. La característica destacable de estos componentes es que son dispositivos seriales, por lo tanto las comunicaciones deben ser realizadas a través del protocolo serial. En consecuencia el software, que se está desarrollando para funcionar en el embebido, deberá emplear un driver que entienda dicha comunicación.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La placa Blue Pill puede ser programada de dos formas distintas: la primera utilizando programas Arduino, empleando el lenguaje Wiring. Este lenguaje anexa una compleja capa de abstracción que afecta la performance del sistema. Por ese motivo se determinó que era conveniente emplear la segunda metodología, que consiste en realizar la programación de estas placas utilizando el lenguaje C y C++, y emplear las bibliotecas HAL y LL [3], que ofrecen la familia de los microcontroladores STM32, las cuales permiten un mayor control de los recursos del hardware y son autogeneradas por la herramienta STM32Cube-MX, que permite configurar la Blue Pill fácilmente. También el código fuente que autogenera el STM32Cube-MX, es creado en lenguaje C únicamente, por lo que por defecto, los proyectos deben ser construidos utilizando ese lenguaje de programación. No obstante, debido a los beneficios que ofrece la programación orientada a objetos con C++ [4] [5], se determinó conveniente programar utilizando también este último lenguaje. Por consiguiente el código del programa embebido está siendo desarrollado en forma hibrida, parte de la configuración inicial de la placa es realizada en C, a través de las bibliotecas HAL, y por otro lado el resto del programa embebido es desarrollado en C++. Adaptando el código fuente que autogenera el STM32Cube-MX y las opciones del compilador arm-none-eabi, exclusivo para la arquitectura ARM Cortex [6], para que el código fuente pueda ser desarrollado utilizando C y C++.

Diseño del Sistema

El diseño de los módulos del programa que se ejecutará en el dispositivo IoT como se muestra en la **Fig. 1** . El sistema es dirigido por un Controlador Principal, el cuál determina las acciones que deberán realizar los distintos componentes de hardware que estarán conectados a la Blue Pill. A su vez como una acción de un componente hardware conlleva distintas tareas complejas, ejecutando un conjunto de instrucciones, se determinó conveniente que estas sean realizadas por los demás Controladores para mayor simplicidad. Para el manejo de los dispositivos seriales se creó el módulo *Serial* encargado de dicha tarea.

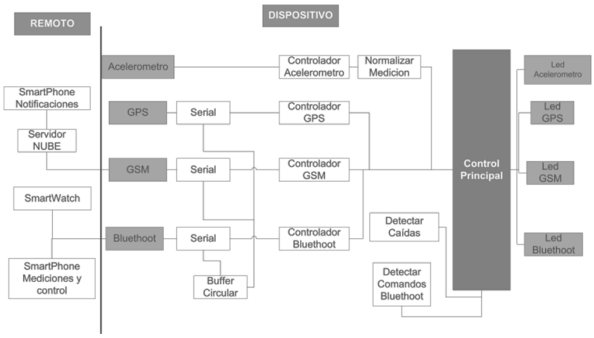


Fig. 2 *Diseño de los módulos que conforman el dispositivo IoT*

Adaptación de bibliotecas para el Driver Serial

El desarrollo del software necesario para poder comunicar de manera eficiente el Bluetooth, GPS y GSM/GPRS con la placa STM32., presenta tres mecanismos distintos para leer o escribir datos. Estas operaciones se realizan utilizando determinados registros de acuerdo al tipo de periférico con que se esté trabajando. El primer mecanismo consiste en la lectura y escritura de forma bloqueante por medio de la técnica de polling, el segundo en el uso de Interrupciones y tercer método utiliza transferencia de datos a través de DMA [7]. Haciendo uso de estas técnicas se desarrolló un driver genérico, que permite realizar la transferencia de datos de forma serial. Esto se efectuó de esta forma debido a que tanto el módulo de bluetooth, como el de GPS y el de GSM/GPRS realizan la transferencia de datos en forma serial mediante los puertos USART[1]. El dispositivo STM32F103C8T6 presenta tres puertos USART independientes, que son accesibles a través de sus pines de conexión. Para poder manipular los registros "Receive Data Register" (RDR) y "Transmit Data Register" (TDR) de la USART [7], se utilizaron parte de las bibliotecas HAL y LL configuradas con el software STM32Cube-MX, que debieron ser adaptadas en su funcionamiento, complementándolas por medio de la creación del módulo Serial (**Fig. 2**) para poder realizar la lectura y escritura continua de datos a través de la USART. Esto se debió a que las funciones de HAL y LL no están adaptadas para realizar dichas operaciones en forma ininterrumpida con la Blue Pill. Por lo que no se lograba completar la totalidad de las funcionalidades que estas otorgan. El módulo serial fue creado de manera que pueda ser configurado dinámicamente, permitiendo seleccionar el mecanismo de operación con el que se desea trabajar los puertos series: ya sea mediante polling, interrupciones o DMA. Una vez escogido el tipo de operación el driver realiza la transferencia de datos desde la USART al dispositivo a través de instrucciones que envían datos al registro TDR del puerto serie. Dado el tamaño limitado del registro, no es inmediato el envío de los bytes de las tramas al dispositivo serial destino. Por ese motivo fue necesario crear un buffer circular de transmisión que amortiguase los datos que se envían al dispositivo serial, hasta que el TDR se encuentre disponible para poder enviar esos bytes de información y otro buffer circular de recepción, con la finalidad de evitar pérdida de información en los datos que envían los dispositivos seriales a la Blue Pill. Estos datos son recibidos en su registro RDR.

El módulo Serial desarrollado está compuesto por una biblioteca que contiene una Clase denominada *Serial,* cuyos métodos pueden ser invocados desde cualquier parte del código fuente del programa embebido con tan solo crear un objeto de dicha clase. Por consiguiente, al iniciar el Controlador Principal se debieron instanciar tres objetos Seriales: el primero de ellos para emplear el GSM, el segundo para el Bluetooth y el tercero para poder acceder al GPS. Luego estos tres objetos son utilizados internamente por los módulos Controladores para acceder a los respectivos dispositivos. Los

[1] Universal Receptor Transmisor Sincrónico/Asincrónico

métodos más importantes de la Clase Serial son: *read y send*. De acuerdo al método de operación con el cual haya sido creado el objeto, el método *read* recibe los bytes que son enviados al puerto serial correspondiente y los retorna a su invocador. A su vez este método puede ser ejecutado de forma bloqueante o no, leyendo los bytes que fueron almacenados en el buffer circular de recepción correspondiente a dicho a objeto. Por otro lado el método *send*, permite enviar cadenas de bytes al puerto serial correspondiente, también de forma bloqueante o no bloqueante. En este último caso, cada vez que sea invocado el método *send*, se irán apilando en un buffer circular las cadenas que se desean enviar, que serán luego transmitidas cuando se libere el registro TDR por medio de un manejador de interrupciones. Para poder enviar varios parámetros al puerto serial, el método *send* utiliza las macros *va_start* y *va_end*, que son utilizadas en ANSI C para acceder a múltiples parámetros. Además para evitar problemas de sincronismo con los buffers circulares, fue necesario deshabilitar y habilitar las interrupciones en ciertas secciones del código fuente del programa. Esto se debe a que la programación del sistema embebido está siendo realizada sin multiprogramación, ejecutándose todo el programa en un único hilo de ejecución.

Utilización del Driver Serial

Los módulos de hardware GSM SIM800L y el Bluetooth HM-10 son dispositivos que deben emplear el módulo serial. Dentro del código fuente de sus Controladores se realiza los envíos de datos y comandos al componente, invocando para ello el método *send* del objeto que se encuentra asociado al mismo. Mientras que por otro lado la recepción de datos y comandos recibidos por estos componentes, es realizado a través del método *read*, para posteriormente poder analizar la información recibida y en base a ella actuar en consecuencia. Los comandos recibidos y enviados por el Controlador Bluetooth, permiten controlar y monitorear de forma remota el dispositivo IoT a través de un Smartphone, durante la etapa de desarrollo. En las etapas subsiguientes se pretende adaptar este Controlador para que pueda conectarse a otros sensores Bluetooth. Por otro lado el SIM800L funciona mediante el envío y ejecución de secuencias de comandos AT. Este componente, actúa con un patrón de ejecución-respuesta, el SIM800L ejecuta un comando AT y retorna una respuesta de confirmación de ejecución correcta. Por ese motivo la ejecución del Controlador GSM no puede ser bloqueante, aplicando las funciones HAL_Delay, debido a que se produciría una espera activa que afectaría el desempeño del sistema. Para resolver esto, muchas aplicaciones de STM32 hacen uso de paralelismo empleando un S.O. de Tiempo Real [8] [9]. No obstante estas soluciones no son óptimas en cuanto a rendimiento, debido a que ocupan demasiado espacio en memoria, entre otros factores. Por lo cual se desarrolló una solución no bloqueante, que se pueda ejecutar en un único hilo de ejecución, aplicando retardos no bloqueantes a través de contadores de tiempos de reloj, por medio de la función HAL_GetTicks. Aplicando dicho concepto, se desarrolló el funcionamiento del algoritmo que ejecutan los distintos comandos en los distintos Controladores. Para esto, se implementaron la codificación de diagramas de estados para la ejecución de los distintos comandos que conforman una operación (**Fig. 3**). Un ejemplo de ello es el caso del ControladorGSM, que para poder enviar un SMS a un número telefónico, debe ejecutar una de secuencia de tres comandos AT. En la figura **Fig. 3** se muestra el algoritmo básico que realiza dicha operación. Al ser no bloqueante, este método debe ser

ejecutado constantemente dentro del bucle del Controlador Principal.

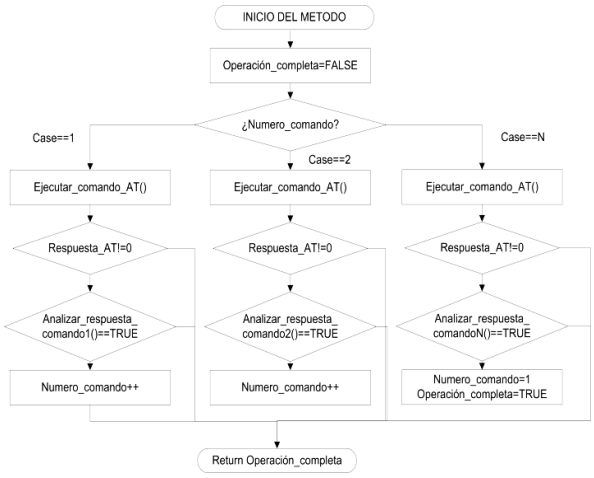


Fig. 3 Diagrama de flujo del Algoritmo básico para ejecutar un comando AT en forma no bloqueante

Un aspecto importante, es que dentro de la función Ejecutar_comando_AT se invoca al método *send,* para enviarle al SIM800L el comando AT que deberá ejecutar. Posteriormente en dicha sección de código, se realiza una espera sin detener el transcurso normal del programa, para poder recibir la respuesta del comando AT previamente ejecutado, utilizando el método *read*.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se ha conseguido desarrollar un módulo que permite realizar una comunicación serial con distintos dispositivos que emplean dicho protocolo, adaptando las bibliotecas HAL y LL, que genera el STM32Cube-MX, De esta forma se han desarrollado algoritmos que permiten controlar y monitorear la ejecución del dispositivo IoT, durante la etapa de desarrollo en forma remota por medio del Bluetooth. Al mismo tiempo se han creado APIS que permiten ejecutar funcionalidades del módulo GSM como: Realizar llamadas telefónicas, enviar SMS y mensajes POST a un servidor, de forma no bloqueante en un único hilo de ejecución.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación dentro del Departamento de Ingeniería e Investigaciones Tecnológicas forma parte del trabajo que dos investigadores se encuentran realizando para su tesis de maestrías. Completan el grupo de investigación cuatro alumnos de Ingeniería que se encuentran finalizando su formación de grado y realizan su iniciación a la investigación.

5. BIBLIOGRAFÍA

[1] G. De Luca, E. Carnuccio, G. Garcia, D. Guilianelli, M. Volker, W. Valiente, V. Raul y M. Vittorio, «Dispositivo de asistencia de personas mediante monitoreo IoT,» de *Libro de Actas Wicc 2018*, Abril 2018, pp. 854-858.

[2] STMicroelectronics, «Specifications STM32F103C8,» [En línea]. Available: https://www.st.com/en/microcontrollers/stm32f103c8.html.

[3] STMicroelectronics, Description of STM32F1 HAL and Low-layer drivers, 2017.

[4] S. Dan, «C++ for Embedded C Programers,» 2013. [En línea]. Available: http://www.dansaks.com/talks/ESC-205.pdf.

[5] upwork, 2017. [En línea]. Available: https://www.upwork.com/hiring/development/c-vs-c-plus-plus/.

[6] ARM, ARM Compiler User Guide, vol. Version 6.7, Cambridge, 2017.

[7] STMicroelectronics, de *STM32F101xx, STM32F102xx,STM32F103xx, STM32F105xx and STM32F107xx advanced Arm®-based 32-bit MCUs*, 2018.

[8] N. Askari, 2018. [En línea]. Available: https://github.com/nimaltd/Sim800_V2.

[9] T. Majerle, 2018. [En línea]. Available: https://github.com/MaJerle/GSM_AT_Lib.

ANÁLISIS DE RENDIMIENTO Y EQUIDAD EN TCP

Claudia N. González, Diego R. Rodríguez Herlein, Carlos A. Talay, Franco A. Trinidad y María L. Almada

Campus Universitario – Oficina 18/ Dpto. Ciencias Exactas e Informática UARG / UNPA
{cgonzalez, dherlein, ctalay}@uarg.unpa.edu.ar, tfrancoalejandro@gmail.com, mluzalmada@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

En una red híbrida la asignación de recursos juega un papel importante, ya que se busca que en las mismas sean compartidos equitativamente. El presente trabajo, muestra las líneas de investigación que se vienen desarrollando sobre la equidad de los algoritmos de control de congestión de TCP aplicadas en sus distintas variantes.

Para este análisis se realizó una serie de ensayos utilizando el simulador NS-2 sobre un escenario híbrido, donde se observa la competencia por los recursos de la red. Para evaluar el rendimiento y el nivel de equidad, se utilizaron como referencia las métricas del thoughput instantáneo, índice de equidad y tiempo de convergencia.

Palabras clave: TCP, red híbrida, ancho de banda, equidad

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 en 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 cuenta 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.

1. INTRODUCCIÓN

El protocolo TCP [1] permite la comunicación orientada a la conexión y el intercambio de datos, es decir que garantiza que los paquetes sean entregados en el mismo orden en el que fueron enviados. Para ello, se le da un número de orden a cada paquete y se espera que el receptor mande un ACK, para garantizar que lo recibió.

Uno de los aspectos más importantes del protocolo TCP, es el manejo de ciertos algoritmos utilizados para realizar el control de congestión, basado en ventana deslizante [2], pudiendo así estudiar la prevención del colapso de la red y de esta forma poder alcanzar un alto grado de rendimiento en la red.

En redes inalámbricas muchos nodos compiten por los recursos disponibles, creando un desafío en la asignación de recursos. Es aquí donde juegan un papel importante el rendimiento y la justicia [3], donde el primero refleja la utilización de la red inalámbrica y la segunda el equilibrio en el reparto de recursos.

En este trabajo como así también en el proyecto de investigación se pretende analizar las métricas a ser consideradas en la asignación

del ancho de banda, utilizando un modelo simple, con dos flujos de datos y considerando las variantes más conocidas del protocolo TCP [4], bajo un escenario híbrido.

Esto permitiría ver cómo es utilizado el ancho de banda y cómo se comportan ante la competencia por este recurso.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El protocolo TCP es el encargado de transportar la mayor parte del tráfico de internet, siendo ésta la causal del rendimiento y funcionamiento del mismo. Es por esto que analizaremos las métricas que deben ser consideradas en la evaluación de los mecanismos de control de congestión.

Se habla de equidad cuando flujos que comparten un enlace, se reparten entre todos por igual el ancho de banda disponible. Se puede ver, de esta manera, la forma, en que un flujo TCP afecta al resto de los flujos, y de cómo es afectado por ellos en términos del throughput.

El concepto de equidad se puede analizar en estas tres categorías:

- Equidad Intra-protocolo (Intra-protocol fairness): el cual hace referencia a la distribución de recursos entre los flujos que ejecutan el mismo algoritmo de control de la congestión, en el mismo entorno de red [5].

- Equidad Inter-protocolo (Inter-protocol fairness): el cual hace referencia a la imparcialidad de la distribución entre flujos que ejecutan diferentes algoritmos en el mismo entorno [6].

- Equidad RTT (RTT-fairness): el cual hace referencia a la distribución de recursos entre flujos que comparten el mismo vínculo de cuello de botella pero que tienen diferentes valores de RTT [7].

A continuación se desarrollan las métricas que se van a utilizar para evaluar cuantitativamente, por ejemplo:

Índice de equidad

En esta línea, Jain [5] propuso un parámetro que permite medir el reparto equitativo de los recursos de red, el cual se define de la siguiente manera:

$$f(x_1, x_2, x_3 \ldots x_n) = \frac{(\sum_1^n x_i)2}{n \times \sum_1^n x_i^2}, \quad 0 \leq f \leq 1$$

Para un conjunto dado de flujos TCP (x_1; x_2; x_3; x_4; x_5… X_n;), esta fórmula calcula el índice de equidad.

El índice de equidad es un valor entre 0 y 1, donde 1, muestra la distribución más equitativa del ancho de banda disponible del enlace entre los flujos de la competencia en una red y 0 muestra la distribución injusta del ancho de banda disponible del enlace.

Tiempo de Convergencia

El tiempo de convergencia de los mecanismos de control de congestión se calcula utilizando la siguiente ecuación

$$T = T_3 - T_2$$

Donde:

T2: Es el tiempo de transmisión inicial del flujo 2.

T3: Es el tiempo en segundos, cuando el valor del throughput del flujo 2 alcanza el 80% del tamaño del throughput del flujo1 [8].

3. RESULTADOS OBTENIDOS/ESPERADOS

El modelo implementado en el simulador de eventos discretos NS-2 (Network Simulator 2, v2.35) muestra un escenario mixto simple con 2 nodos y sus enlaces cableados, una estación base y dos nodos con sus enlaces inalámbricos,

como se observa en la siguiente figura (Figura 1).

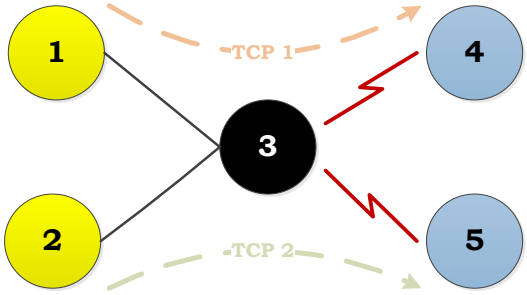


Figura 1. Modelo de estudio

Se realizaron ensayos donde se confrontaron 15 protocolos con sus respectivos controles de congestión, de las cuales los más relevantes para nuestro análisis fueron Reno, Compound, Cubic, Vegas, Illinois.

De cada una de estas simulaciones se obtuvo el throughput promedio, mediante el uso de scripts AWK, para filtrar los valores correspondientes del archivo de traza generado por NS2. Luego se utilizó gnuplot para graficar los resultados.

Para la comparación de las distintas variantes de TCP en competencia se utilizó una métrica resultante del producto del throughput instantáneo, índice de equidad y el tiempo de convergencia definido anteriormente.

A continuación se muestran, gráficos obtenidos sobre el modelo escogido (Figura 1), de cómo será el comportamiento cuando dos protocolos distintos compiten por los recursos, sobre un canal común llevando a un nivel estable de transmisión de datos.

Considerando que tenemos dos protocolos A y B, donde inicia primero el protocolo A y al cabo de un tiempo lo sigue el protocolo B.

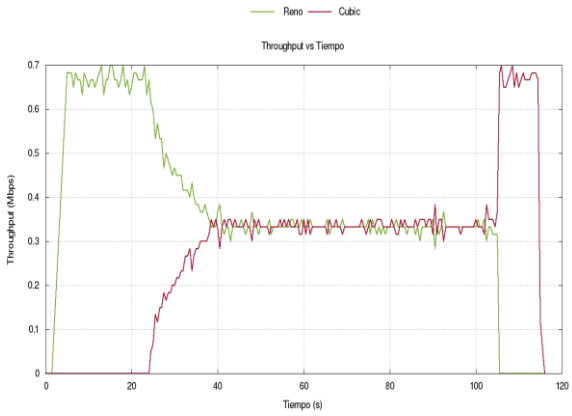


Figura 2. Throughput - Reno vs. Cubic

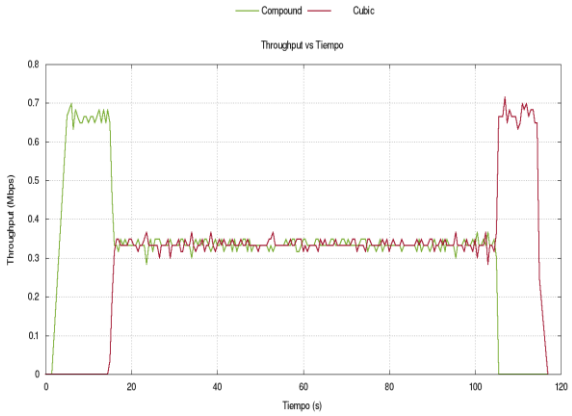


Figura 3. Throughput - Compound vs. Cubic

En la figura 2 y 3 se puede observar cómo es la evolución de thoughput instantáneo, cuando estando un protocolo A disponiendo del total del ancho de banda se introduce a la transmisión el protocolo B, y cómo se llega a una distribución equitativa del ancho de banda.

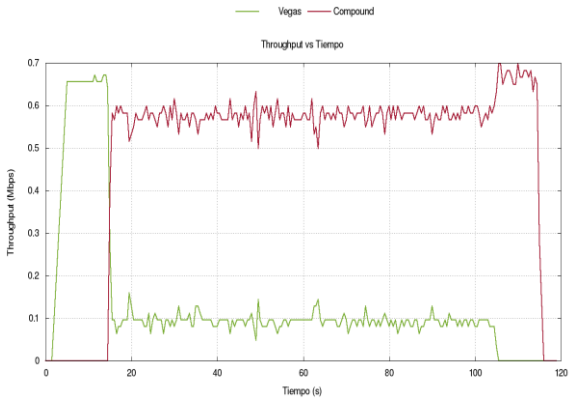


Figura 4. Throughput - Vegas vs. Compound

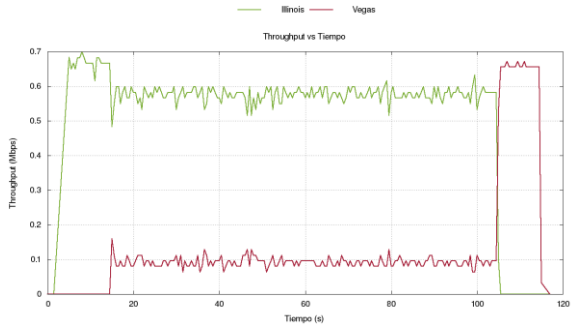


Figura 5. Throughput - Illinois Vs. Vegas

En la figura 4 y 5 se puede observar cómo la evolución del thoughput no es equitativa, ya que en ambos casos aunque Vegas inicia su transmisión primera o comienza en segundo término, al desarrollarse la contienda, Vegas resigna su participación en el uso de ancho de banda del canal.

Observando las figuras 2 y 3, además de lograr en ambos casos un reparto justo del ancho de banda, se puede obtener un primer análisis y ver que los tiempos de convergencia son distintos. Se puede observar que mientras que en la figura 2 le lleva un tiempo determinado llegar a la posición de equilibrio, en la figura 3 ese tiempo es menor.

Actualmente en el proyecto de investigación estamos trabajando en el análisis de cómo aplicar las métricas a casos descriptos en las Fig.1, Fig.2, Fig. 3 y Fig. 4 para lograr una interpretación más acabada del fenómeno de equidad.

Como así también, ampliar el estudio con ensayos donde se modifique el modelo con distintos RTT, para realizar las comparaciones con las métricas RTT-fairness.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación 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 integran 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.

5. 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] Montuno, K., Zhacfi, Y. “Fairness of Resource Allocation in Cellular Networks: A Survey”, In: Resource Allocation in Next Generation Wireless Networks pp. 249–266 (2006).

[4] 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.

[5] R. Jain, D.M. Chiu, and W. Hawe, “A Quantitative Measure of Fairness and Discrimination for Resource Allocation in Shared Systems”, DEC TR-301, Littleton, MA: Digital Equipment Corporation, 1984.

[6] Floyd, S., and Fall, K., "Promoting the Use of End-to-End Congestion Control in the Internet", IEEE/ACM Transactions on Networking, August 1999.

[7] Lisong Xu, Khaled Harfoush, and Injong Rhee, “Binary Increase Congestion Control for Fast Long-Distance Networks”, INFOCOM 2004.

[8] D. Chiu and R. Jain. “Analysis of the Increase/Decrease Algorithms for Congestion Avoidance in Computer Networks.” Journal of Computer Networks and ISDN, 17(1):1 14, 19

Despliegue de una Red LoRaWAN en la Universidad Nacional de Misiones

Eduardo O. Sosa, Milton E. Sosa, Diego A. Godoy

Secretaría de Investigación y Posgrado (SECIP). Facultad de Ciencias Exactas, Químicas y Naturales; Félix de Azara 1552 - N3300LQH - Posadas

Universidad Nacional de Misiones (UNaM)

{es; mesosa; dgodoy}@fceqyn.unam.edu.ar

Resumen

Existiendo en el mundo más de siete mil millones de dispositivos conectados, y un crecimiento esperado sin precedentes para los próximos años; Internet de las Cosas (IoT) es uno de los temas más candentes de la actualidad. El control de entornos físicos en un futuro próximo requiere nuevas tecnologías inalámbricas que puedan extender la vida útil de la batería y cubrir áreas extensas. LoRaWAN es una de las tecnologías más destacadas que cumplen con estos requisitos atrayendo la atención tanto de la academia como de la industria. En este trabajo se presenta el desarrollo, configuración e implementación de dos nodos LoRaWAN como puertas de enlace a los servidores de The Things Network. Este servicio es prestado a la ciudadanía de la ciudad de Posadas en forma abierta, y soportado oficialmente por la red de datos de la Universidad Nacional de Misiones (UNaM). Dichos *gateways* son capaces de dar soporte oficial y privado a situaciones críticas, de emergencia o desastres naturales, como también a la investigación y desarrollo.

Palabras clave: Internet de las Cosas (IoT), LoRa, LoRaWAN, Redes de Sensores

Contexto

El proyecto se relaciona con otros sobre aplicaciones de conceptos de Internet de las Cosas (IoT), Internet del Futuro y sistemas inteligentes. Es el producto de desarrollos en el marco de la Secretaría de Investigación y Posgrado de la Facultad de Ciencias Exactas, Químicas y Naturales de la Universidad Nacional de Misiones. En este caso el proyecto aprobado ha sido "Localización de Objetos en Internet de las Cosas", nomenclado como 16Q652-PI de grupo consolidado, habiendo sido aprobado y dado de alta por Resolución 291/18.

Se han coordinado actividades y consultas al Centro de Investigación sobre Tecnologías de la Información y Comunicación (CITIC) de la Universidad Gastón Dachary (UGD) de manera de avanzar en el tema realizando las consultas pertinentes a otros involucrados en la cuestión, considerando que es una temática considerada en esa casa de estudios.

Introducción

Las redes móviles celulares actuales se diseñaron principalmente para interacciones de tipo persona-persona y persona-máquina enfocadas en aplicaciones/servicios específicos tales como telefonía, intercambio de SMS/MMS, descargas multimedia y transmisión por secuencias. Esto no se condice con la visión y evolución de la Internet de las cosas (IoT), que requiere conectividad también para dispositivos desatendidos. Esto define de facto un nuevo paradigma de comunicación (y tráfico) que involucra poca o ninguna interacción humana, y por lo tanto, a menudo se lo conoce como comunicaciones "máquina-máquina"; ó "tipo

máquina”. Considerando la necesidad de cobertura espacial, de tiempo de funcionamiento y tipos de datos a transferir, dos tecnologías cumplen con los requerimientos: Redes celulares y Redes de área amplia de baja potencia (LPWAN).

El Proyecto de Asociación de Tercera Generación (3GPP) ha lanzado varias actividades para mejorar los estándares clásicos de las comunicaciones M2M, incluyendo la Cobertura Extendida-GSM (EC-GSM), la Comunicación Tipo Máquina (LTE-M) mejorada de LTE y/o la banda estrecha IoT (NB-IoT). Sin embargo, las tecnologías antes mencionadas están tomando un tiempo para la estandarización, no alcanzando aún la madurez y difusión completas. Dichos "retrasos" en el tiempo de comercialización de los estándares celulares compatibles con IoT, abrieron el camino a los operadores de redes especializados que venden conectividad IoT a través de tecnologías inalámbricas de largo alcance y baja potencia como SigFox [1] LoraWAN [2], Weightless [3]] e Ingenu [4].

LoRaWAN es un estándar inalámbrico de LoRa . La arquitectura de referencia LoRaWAN, es mostrada en la Figura 1.

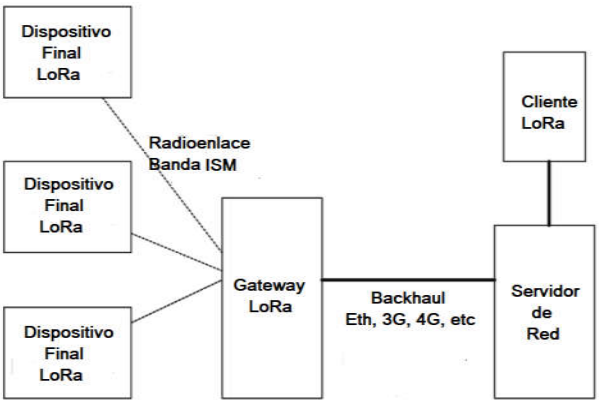


Figura 1. Arquitectura LoRaWan

Los dispositivos finales detectan y transmiten datos de campo hacia las puertas de enlace puertas de enlace que recopilan los datos de los nodos finales y los envían a un servidor de red donde se encuentra toda la "inteligencia" de la red. Los nodos finales no están asociados a ninguna puerta de enlace específica, sino que transmiten sus mensajes que son recibidos y reenviados por cualquier puerta de enlace a su alcance

El servidor de red se encarga de eliminar los duplicados en los mensajes de enlace ascendente, ejecutar la lógica MAC y administrar los parámetros de RF de LoRa.

LoRaWAN opera en el espectro de radio Sub-Gigahertz sin necesidad de licencia (bandas Industrial, Científica y Médica –ISM-), con frecuencias de portadora específicas para cada región del planeta. Los nodos instalados en la UNaM, han sido configurados en la banda de 915 MHz, en un todo de acuerdo a las bandas asignadas por continentes.

Para los nodos desplegados se han adoptado las frecuencias LoRaWAN para los Estados Unidos de América definiendo 64 canales de 125 kHz de 902.3 a 914.9 MHz. Ocho canales de enlace ascendente de 500 KHz en incrementos de 1,6 MHz de 903 MHz a 914 MHz. Los ocho canales de enlace descendente tienen un ancho de 500 kHz desde 923.3 MHz hasta 927.5 MHz. La potencia de salida máxima es de +30 dBM.

El corazón de LoraWAN es LoRa, una técnica de modulación patentada de espectro expandido, basada en un derivado de Chirp Spread spectrum (CSS) [5]. CSS es una técnica de espectro ensanchado que utiliza pulsos modulados en frecuencia lineal de banda ancha para codificar información, como ejemplo, una señal se transmite en ráfagas cortas, "saltando" entre frecuencias en una secuencia pseudoaleatoria

Hemos basado nuestro proyecto en LoRaWAN, debido a el rendimiento de éstas redes en términos de cobertura, latencia de extremo a extremo y tasa de extracción de datos (DER), como también del diseño de la red, y otros de configuración de la capa física que caracterizan el segmento inalámbrico (factor de expansión, tasa de codificación de protección, ancho de banda del canal, potencia emitida).

Líneas de Investigación

La investigación académica y científica sobre LoRaWAN está en su génesis y, hasta el momento, está dirigida principalmente a la evaluación del rendimiento de las capas física y de acceso al medio del estándar de referencia OSI [6].

En el seno de la Facultad de Ciencias Exactas, Químicas y Naturales se desarrollan actividades en proyectos de investigación que presentan un ordenamiento disciplinar más acotado a procesos constitutivos del campo de la Tecnologías de la Información y Comunicaciones, con una lógica de trabajo que ha nutrido de experiencias, prácticas y saberes centrales al efecto de dar cuenta de las mejores formas de avanzar sobre temas del dominio especificado. Por ello los estamentos involucrados proponen el desarrollo de líneas de investigación que, transversalmente, den cuenta de los objetivos perseguidos originando compilación de ideas proyecto, desarrollos, captura de información, conceptualización, análisis y producción de elementos utilizables en escenarios reales, con posibilidades ciertas de extrapolación trabajos del futuro.

Hemos establecido una sucesión continua y ordenadora de actividades de estudio, reflexiones sistemáticas y creativas, tormentas de ideas y discusiones alrededor del tema Internet de las Cosas, enlazando entre si uno o varios proyectos de trabajo para así desarrollar actividades académicas e intelectuales dinámicas, en medio de búsquedas, hipótesis, logros, metas y objetivos para producir, construir y aumentar conocimientos sobre redes sub-gigahertz. De esta manera actuamos colectivamente a situaciones que demandan procesos sistemáticos fundamentalmente interdisciplinarios.

Materiales, Métodos e Implementación

Las puertas de enlace forman el puente entre los dispositivos y los servidores finales. Los dispositivos usan redes de bajo consumo como LoRaWAN para conectarse a la Puerta de enlace, mientras que la Puerta de enlace usa la red TCP/IP de la Universidad para conectarse a la Red de Cosas. Todas las puertas de enlace al alcance de un dispositivo recibirán los mensajes del dispositivo y los reenviarán al servidor. En destino, se eliminarán los mensajes duplicados y se seleccionará la mejor puerta de enlace para reenviar los mensajes en cola por el enlace descendente. Una única puerta de enlace puede servir a miles de dispositivos.

La necesidad de desplegar *gateways* en el país, es obvia toda vez que se observa la dispersión de nodos públicos accesibles a lo largo del país (Figura 2).

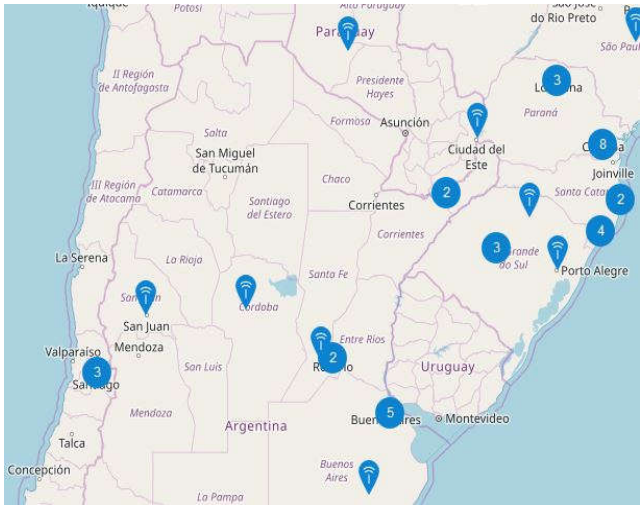


Figura 2. Nodos desplegados en el norte Argentino

Las puertas de enlace son enrutadores equipados con un concentrador LoRa, lo que les permite recibir paquetes LoRa. Normalmente se encuentran dos tipos de puertas de enlace: a) aquellas que se ejecutan en un firmware mínimo, y solo ejecutan el software de reenvío de paquetes y b) las que ejecutan un sistema operativo, para el cual el software de reenvío de paquetes se ejecuta como un programa en

segundo plano. Nuestros nodos se encuadran entre los primeros.

Como puerta de enlace LoRa hemos utilizado un módulo RAK831 [7] capaz de recibir en diferentes canales de frecuencia al mismo tiempo y de demodular la señal LoRa sin conocer el factor de propagación utilizado del nodo emisor.

Se ha utilizado un host Raspberry Pi 3 para su correcto funcionamiento, conectándose al módulo RAK831 a través de SPI[1].

En la Figura 3 puede observarse la zona abarcada por los nodos instalados en la Facultad de Ciencias Exactas tanto en la sede centro. Como en el anexo de bioquímica y farmacia.

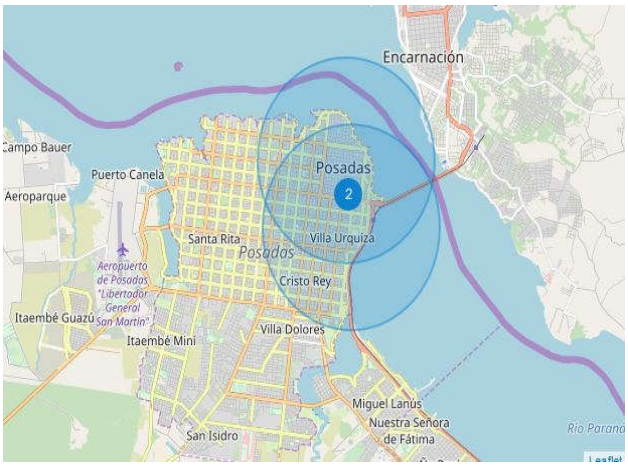


Figura 3. Zona cubierta por los nodos en la FCEQyN

Haciendo uso de las herramientas disponibles para nodos IoT, se han configurado sensores de temperatura, humedad y temperatura de en cada una de las cajas estanca, con los servicios de Thingspeak [8], Figuras 4 a, b y c.

Formación de Recursos Humanos

Los integrantes de grupo de Investigación, son investigadores formados, que forman un grupo consolidado, habiendo obtenido el primer lugar en el órden de mérito entre todos los proyectos de investigación presentados en la Universidad Nacional de Misiones, habiendo sido evaluado externamente.

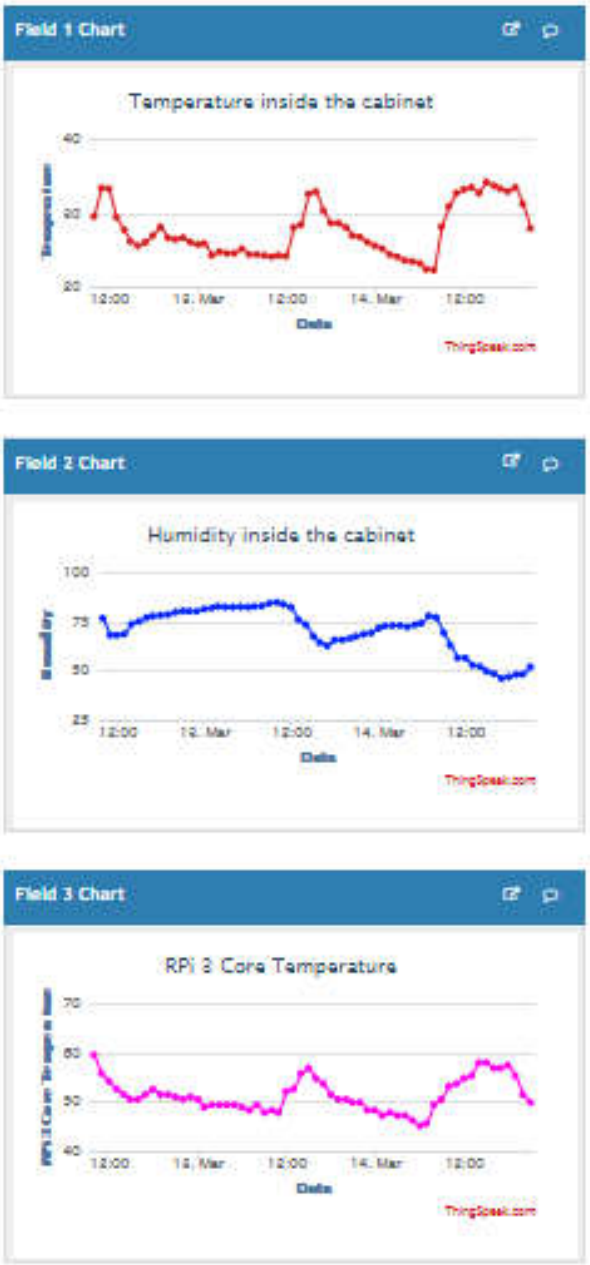


Figura 4. Parámetros físicos estancos de gateway de nodo anexo, a) temperatura, b) Humedad c) T Core Pi

[1] **SPI** es un acrónimo para referirse al protocolo de comunicación serial Serial Peripherical Interface

Referencias

[1] Sigfox, "Sigfox, the world's leading IoT services provider.," 03 2019. [Online]. Available: https://www.sigfox.com/en.

[2] LoRa Alliance, «What is the LoRaWAN Specification?,» Marzo 2019. [En línea]. Available: https://goo.gl/wqdAnP.

[3] Weightless SIG, «Weightless. Simply better.,» [En línea]. Available: http://www.weightless.org.

[4] I. Inc., "IOT Connectivity that works," Marzo 2019. [Online]. Available: https://www.ingenu.com.

[5] B. Reynders and S. Pollin, "Chirp spread spectrum as a modulation technique for long range communication," in *2016 Symposium on Communications and Vehicular Technologies (SCVT)*, Mons, Belgium, 2016.

[6] International Organization for Standardization, «Estándar ISO 7498-1,» ISO/IEC, Ginebra, 1994.

[7] Shenzhen Ruike Huilian Technology Co., Ltd., «RAk Products,» [En línea]. Available: https://goo.gl/vvwWZe. [Último acceso: marzo 2019].

[8] The MathWorks, Inc., «Understand Your Things,» [En línea]. Available: https://thingspeak.com/channels/557580. [Último acceso: Marzo 2019].

“Determinación de la eficiencia y Estrategias de Tolerancia a Fallos en Arquitecturas Multiprocesador para aplicaciones de procesamiento de datos”

Jorge R. Osio[1,2], Diego Montezanti[1,4],Eduardo Kunysz[1],Daniel Martin Morales[1,3]

[1]Instituto de Ingeniería y Agronomía - UNAJ
[2]UIDET CeTAD –Fac. de Ingeniería -UNLP
[3]Codiseño HW SW para Aplicaciones en Tiempo Real - UTN - FRLP
[4]Instituto 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 de procesamiento sobre las imágenes médicas. Por otra parte, debido a que el procesamiento en paralelo requiere de la implementación de sistemas de múltiples procesadores en dispositivos reconfigurables, se ha incorporado al proyecto una línea en la que se está trabajando en el desarrollo de metodología de tolerancia a fallos transitorios, que son cada vez más frecuentes en las arquitecturas paralelas, y que afectan especialmente a las aplicaciones de cómputo intensivo y ejecuciones de larga duración.

Palabras clave: *arquitecturas paralelas, procesamiento de imágenes, checkpoints de capa de sistema, tolerancia a fallos, sistemas multicores, dispositivos reconfigurables.*

Contexto

Las líneas de Investigación descritas 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).

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 con 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 Instituto LIDI.

1. Introducción

En los últimos años, se ha buscado expandir el concepto del procesamiento paralelo con computadoras basadas en multicores hacia la utilización de plataformas de procesamiento más específicas. Para obtener mayor eficiencia, los fabricantes de computadoras de altas prestaciones, han introducido unidades FPGAs(arreglo de compuertas programables en campo) en su diseño como soporte para el cómputo ([1-3]).

Si bien el estudio de sistemas paralelos con múltiples procesadores, 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 de SW y Concurrencia en VHDL (lenguaje de descripción de hardware). Adicionalmente, se investiga sobre la tolerancia a fallos de sistemas con gran poder de cómputo y basados en HPC.

Plataformas FPGAs para procesamiento paralelo

La implementación de 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 HW permite combinar las FPGAs con el paralelismo obtenido mediante sistemas multicore para alcanzar la máxima eficiencia.

Entre las ventajas que proveen las FPGAs actuales para el cómputo paralelo se pueden enumerar:

- SoftCores (o procesadores embebidos) que permiten realizar las tareas de administración de datos de baja tasa. Y la posible conexión de IP-Cores (CoProcesadores) específicos con las siguientes características para realizar el procesamiento duro de señales:
- Capacidad de procesamiento con millones de MACs por segundo (operaciones de Multiplicación / acumulación)
- DSP o Delay Locked Loops (DLL) que permiten la multiplicación o división de la frecuencia de reloj, entre otras tareas.
- Interfaces DRAM / SRAM de alta velocidad y rendimiento.
- Manejo del ancho de banda en señales y buses ahorrando pines I/O.
- Registros de desplazamiento, útiles para buffers de línea o FIFOs.
- RAM distribuida para almacenar coeficientes o pequeñas FIFOs.
- Block RAM con capacidad "true dual-port" para almacenar datos de fotograma, líneas o porciones de imagen, grandes tablas o FIFOs.

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 Fallo en sistemas multicore

El aumento en la escala de integración, con el objetivo de mejorar las prestaciones en los procesadores, y el crecimiento del tamaño de los sistemas de cómputo, han producido que la confiabilidad se haya vuelto un aspecto crítico. En particular, la creciente vulnerabilidad a los fallos transitorios se ha vuelto altamente relevante, a causa de la capacidad de estos fallos de alterar los resultados de las aplicaciones.

El impacto de los fallos transitorios aumenta notoriamente en el contexto del Cómputo de Altas Prestaciones, debido a que el Tiempo Medio Entre Fallos (MTBF) del sistema

disminuye al incrementarse el número de procesadores. En un escenario típico, en el cual cientos o miles de núcleos de procesamiento trabajan en conjunto para ejecutar aplicaciones paralelas, la incidencia de los fallos transitorios crece en el caso de que las ejecuciones tengan una elevada duración, debido a que el tiempo de cómputo y los recursos utilizados desperdiciados resultan mayores [9 - 10]. Por lo tanto, el alto costo 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, el foco está puesto en lograr detección y recuperación automática de los fallos silenciosos, que no son detectados por ninguna capa del software del sistema, por lo que, sin producir una finalización abrupta, son capaces de corromper los resultados 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 [3-6] y de la investigación de la Tolerancia a fallos en sistemas multicore ([3], [7] y [8]).

Las líneas de investigación se enmarcan fundamentalmente en la búsqueda de técnicas innovadoras de procesamiento paralelo, sobre diferentes arquitecturas, que optimicen las prestaciones en equipos de imágenes médicas (PET, X-Ray, Ecógrafo). Adicionalmente, se pretenden encauzar 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, además dirige proyectos de investigación desde el año 2015 en la UNAJ. Actualmente, el grupo de trabajo investiga en los temas: Procesamiento Digital de Imágenes Médicas sobre plataformas FPGA, Procesamiento de Imágenes en sistemas HPRC y tolerancia a fallos pertenecientes a estudios de Maestrías y Doctorados que se realizan en el marco de acuerdos de colaboración entre el Laboratorio CeTAD,el Instituto 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 de Ingeniería Informática y Ciencias de la Computación [3 - 7].

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, procesamiento de imágenes.

Temas de Estudio e Investigación

- Implementación de un sistema multiprocesador en Dispositivos Lógicos Programables (FPGAs).
- Análisis y determinación del desempeño logrado en el procesamiento de imágenes mediante la combinación de cómputo paralelo y concurrencia.
- Diseño e implementación de una estrategia que permite al sistema detectar y recuperarse automáticamente de los errores producidos por fallos transitorios. La implementación está basada en replicación de software y es totalmente distribuida; está diseñada para operar 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 resultados satisfactorios en relación los objetivos principales:

- En cuando al procesamiento de imágenes sobre dispositivos FPGAs, se han implementados varios algoritmos en sistemas con dos microblaze, en donde el acceso a los datos se realizó mediante memoria compartida DDR2-SDRAM de alta velocidad y con múltiples puertos. El procesador master es el encargado de coordinar la lectura y procesamiento de los datos mediante pasaje de mensajes. Respecto a los resultados, el tiempo de procesamiento se disminuye casi al 50% al utilizar dos procesadores. Se han obtenido tiempos muy buenos en los filtros de mediana y de media (4 ms por fila), pero este se duplica al ejecutar algoritmos con mucho cálculo matemático como el algoritmo de sobel (7,5 ms por fila).
- En cuanto a la tolerancia a los fallos transitorios, se ha diseñado e implementado una estrategia de detección basada en la replicación de procesos y el monitoreo de las comunicaciones, y un mecanismo de recuperación basado en el almacenamiento de un conjunto de checkpoints distribuidos de capa de sistema. La recuperación es automática y se realiza regresando atrás al checkpoint correspondiente a la cantidad de veces que se detecta un fallo determinado. Esta estrategia se encuentra en proceso de validación experimental. Para ello, se ha desarrollado un conjunto de casos de prueba. Además, se ha diseñado una estrategia de recuperación alternativa, para el caso en el que se cuente con checkpointing de capa de aplicación. En este caso, es posible almacenar sólo el último checkpoint luego de validarlo.

Para el año en curso, se esperan alcanzar importantes resultados en el área de cómputo paralelo y concurrencia posibilitados por los sistemas basados en FPGAs. Con esta arquitectura, constituida por varios procesadores implementados en una misma FPGA se espera mejorar el desempeño logrado con el sistema multiprocesador basado en master-slave y memoria compartida.

Se pretende medir la eficiencia de ejecución lograda con el sistema multiprocesador mediante los siguientes algoritmos basados en operadores de ventana:

- Filtro de media
- Algoritmo de mediana, erosión y dilatación
- Algoritmos de detección de borde (Sobel)

En la línea de tolerancia a fallos, se ha diseñado e implementado una metodología distribuida basada en replicación de software, diseñada específicamente para aplicaciones paralelas científicas de paso de mensajes, capaz de detectar los fallos transitorios que producirían resultados incorrectos y recuperar de manera automática las ejecuciones [8]. En este tipo de aplicaciones, cuyos procesos cooperan para obtener un resultado, la mayor parte de los datos relevantes son comunicados entre tareas. Por lo tanto, la estrategia de detección consiste en la validación de los contenidos de los mensajes a enviar y de los resultados finales. Esta solución representa un término medio entre un alto nivel de cobertura frente a fallos y la introducción de un bajo *overhead* temporal; además, no se realiza trabajo para detectar fallos que no afectan a los resultados. Por otra parte, de esta forma, la corrupción de los datos que utiliza un proceso de la aplicación se mantiene aislada en el contexto de ese proceso, evitándose la propagación a otros procesos. Así, no solamente se mejora la confiabilidad del sistema, sino que también disminuye la latencia de detección, y por lo tanto el tiempo luego del cual se puede relanzar la aplicación, lo cual es especialmente útil en ejecuciones prolongadas, y beneficia al sistema aprovechando la redundancia intrínseca presente en la arquitectura multicore.

Para recuperar al sistema de los efectos del error, la propuesta se basa en restaurar la aplicación a un estado seguro previo a su ocurrencia. Para lograr este objetivo, se ha integrado la detección con un mecanismo basado en múltiples *checkpoints* coordinados en capa de sistema, construidos con la librería DMTCP (que proporcionan cobertura en el caso de que un *checkpoint* resulte afectado por un fallo, imposibilitando la recuperación desde él),

o la utilización de un único *checkpoint* no coordinado de capa de aplicación (construido ad-hoc, basándose en el conocimiento de la aplicación), que puede ser verificado para asegurar la integridad de sus datos.

Para una aplicación de prueba, se ha desarrollado un modelo formal de verificación, que permite constatar la eficacia del mecanismo de recuperación. Se encuentra en fase de implementación una experimentación basada en casos de inyección controlada de fallos que valide las predicciones dadas por el modelo mencionado. Se busca además caracterizar el comportamiento temporal de cada nivel de solución (sólo detección, recuperación basada en múltiples *checkpoints* de capa de sistema o recuperación basada en un único *checkpoint* de capa de aplicación) y verificar el comportamiento pronosticado.

El uso de estas estrategias posibilitaría prescindir de la utilización de redundancia triple con votación para detectar y recuperar de fallos transitorios, proveyendo opciones que proporcionan flexibilidad para adaptarse a los requerimientos o limitaciones del sistema. Además, como estos fallos no requieren reconfiguración del sistema, la recuperación puede realizarse mediante re-ejecución en el mismo coreen 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 dentro del Área Arquitectura, Sistemas Operativos y Redes.

En este proyecto existe cooperación a nivel nacional. Hay un Doctor, un investigador realizando Doctorado y dos realizando Maestrías en temas relacionados con simulación de sistemas multiprocesador, tolerancia a fallos, sistemas embebidos y sistemas multicore en HPRC.

Adicionalmente, se cuenta con la colaboración de estudiantes avanzados.

5. Bibliografía

[1] O. Mencer, K. Tsoi, S.Craimer, 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]J. Osio, D Montezanti, E. Kunysz, Morales M., "Análisis de eficiencia y tolerancia a fallo en Arquitecturas Multiprocesador para aplicaciones de procesamiento de datos ", UNNE, Corrientes, WICC 2018.

[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", Ushuahia, 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]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

[8] 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.

[9] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing".Pearson Addison Wesley, 2003.

[10] 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.

“Diseño de aplicaciones de IoT para la solución de problemas en el medio socio productivo”

Jorge R. Osio[1,2], Marcelo Cappelletti[1,2] , Gabriela Suarez[1], Leonel Navarro[1], Florencia Ayala[1], Juan. E Salvatore[1], Daniel Alonso[1], Diego Encinas[1,4], Daniel Martín Morales[1,3]

[1]Instituto de Ingeniería y Agronomía - UNAJ
[2]UIDET CeTAD –Fac. de Ingeniería - UNLP
[3]Codiseño HW SW para Aplicaciones en Tiempo Real - UTN - FRLP
[4] Instituto LIDI-Fac. de Informática - UNLP

{josio,mcappelletti, jsalvatore, dalonso, martin.morales,dencinas}@unaj.edu.ar

Resumen

El presente proyecto consiste en la investigación de técnicas de internet de las cosas (IoT) y sus aplicaciones en la resolución de problemas 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, continuar con los temas propuestos en 2018 en relación a los sectores tales como el hortícola y florícola, que carecen de soporte y no disponen de herramientas para automatizar procesos y recibir información crítica en tiempo real. Por otro lado, se continuará las pruebas de funcionamiento de un algoritmo inteligente de eficiencia energética que permita minimizar el consumo en sistemas alimentados con energía convencional y energía alternativa (este último es muy utilizado por el sector productivo frori-hortícola). Se continúa con el sistema de control a distancia mediante IoT del robot explorador. Este año se incorpora la aplicación de internet de las cosas en el relevamiento de información relativa a la lectura de sensores dedicados a medir la contaminación en arroyos. Por último, se comenzará con la prueba piloto del laboratorio remoto propuesto en el proyecto de investigación dirigido por la Dra. María Joselevich.

Palabras clave: *Internet de las cosas (IoT), Eficiencia Energética, Laboratorios Remotos, Contaminación en arroyos, Robótica telemetría, wifi, gsm/gprs, Software Embebido.*

Contexto

Las líneas de Investigación aquí desarrolladas son 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 lleva a cabo en la Universidad Nacional Arturo Jauretche (UNAJ) y es continuación del proyecto presentado en la edición anterior de WICC [1] y [2].

Por otra parte, se tiene financiamiento en el marco del programa "Universidad, Diseño y Desarrollo Productivo" del Ministerio de Educación a través del proyecto "Sistema de eficiencia energética" y del proyecto de investigación aprobado en la convocatoria "UNAJ Investiga 2017". En relación al sistema de detección de contaminación en arroyos se dispone de financiamiento mediante el proyecto de vinculación tecnológica del programa "Universidades agregando valor 2018" en colaboración con el Municipio de Florencio Varela.

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 Hortí-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 sí y la conectividad a internet que provee a las redes de objetos [3].

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, es por eso que esta área está dentro de los objetivos. El factor ambiental es una prioridad, en nuestro caso nos concentramos en la eficiencia energética y la contaminación del agua. Para este tipo de ambientes, donde los sistemas están expuestos a los efectos de la humedad y temperatura, se requiere tomar recaudos adicionales para mantener su integridad y correcto funcionamiento. Además, se dispone de un conjunto de protocolos de comunicación e interfaces como GSM/GPRS, Wifi, bluetooth y zigbee que se pueden seleccionar en función de los servicios de comunicación disponibles en donde se implementará el sistema. Estos protocolos, requieren un sistema de procesamiento basado en Arquitecturas de Sistemas Embebidos como las desarrolladas en [4].

Actualmente, se considera una problemática a tratar de manera urgente el uso eficiente de la energía [5], 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 [6]. 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.

Respecto a la aplicación de IoT en Laboratorios remotos, actualmente se están explotando las TICs para la enseñanza en aulas virtuales, es indispensable para dar soporte a estas plataformas el diseño de sistemas de control que permitan controlar y relevar el estado de los sistemas en estudio de forma remota. Desde el proyecto se diseñó una interfaz de Hardware y Software embebido entre servidor y el laboratorio físico para ser utilizado en el proyecto de la Dra. María joselevich [7].

Metodología

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.

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.

En cuanto a las soluciones tecnológicas en las actividades florihortícolas y de contaminación en arroyos requieren de la obtención de datos de redes de sensores de diferentes tipos, para la obtención de características 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 estará formado por un sistema de procesamiento (basado en tecnologías de microcontroladores, arduino, LPC1769 y raspberry Pi), al cual se conectarán los sensores y el dispositivo de comunicación GSM/GPRS.

Para el sistema de control de robot explorador, se propone utilizar sistemas embebidos de última generación, los cuales poseen alto poder de cómputo y bajo costo (sistema de control de robot basado en raspberry con conectividad wifi y cámara de 8MPx). Se implementarán sobre estos sistemas embebidos los algoritmos de automatismo y control necesarios para ejecución de las acciones del sistema robótico.

Respecto al laboratorio remoto, Se implementó un sistema de control del laboratorio físico basado en raspberry donde se configuró un servidor Apache para el acceso remoto a la cámara y el software motion para el control de la cámara. Adicionalmente, se configuraron los pines I/O de la placa para el control del sistema generador de onda. Para el acceso al laboratorio se configuró un plugin de moodle que permite a los usuarios otorgar turno y es compatible con los sistemas académicos.

Antecedentes del Grupo de Trabajo

Se pretenden encauzar 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. Además, dirige proyectos de investigación desde el año 2014. Actualmente, el grupo de trabajo investiga en los temas: Trasmisión de datos IoT (mediante la interfaz Ethernet y GSM/GPRS), Eficiencia Energética 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],[2] y [8]).

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 de Arquitecturas de Procesadores Embebidos, IoT, sistemas operativos embebidos, protocolos de comunicación TCP y UDP y robótica.

En cuanto a la eficiencia energética, consideramos que las soluciones existentes no han 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 (técnicas de IA).

En cuanto a la investigación aplicada al sector fruti-hortícola, se plantea el diseño de

una red de sensores que interactúan entre sí 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 requieren bajo consumo y exigen el uso energía alternativa. Respecto a la determinación de la contaminación en arroyos las características del sistema son similares, solo cambia que este deberá funcionar en un medio líquido y se deberán investigar los efectos del medio sobre las comunicaciones inalámbricas.

En cuanto a la robótica, el robot explorador está tomando popularidad, principalmente por la necesidad de realizar mantenimiento en sectores altamente contaminantes como por ejemplo en los reactores nucleares. 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 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 ([9] y [10]) 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, telemetría y en el sistema del laboratorio remoto ([5] y [7]). Determinación del efecto del medio líquido sobre las señales inalámbricas para la determinación de la contaminación de arroyos en tiempo real. Implementación de un algoritmo básico de redes neuronales para la optimización del consumo y la toma de decisiones 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 [4]. Durante las pruebas realizadas se obtuvieron excelentes tiempos en la interacción entre servidor y la plataforma de lectura de sensores y control de actuadores.
- Se implementó un sistema de telemetría mediante la interfaz GSM/GPRS con el módulo M92 de Quectel, que permite supervisar el estado del cultivo, emitir alarmas por sobretemperatura e incendio y accionar el riego automático desde una aplicación móvil, el cual fue probado en campo y se obtuvieron excelentes resultados.
- Se instaló en un laboratorio de la UNAJ un sistema que permite optimizar el consumo energético mediante la ejecución de algoritmos que tienen en cuenta diferentes factores (iluminación, presencia, temperatura). En una primera instancia se midió el consumo sin el sistema de control y actualmente se están obteniendo resultados del consumo con el sistema de control instalado. Este sistema posee conexión con un servidor provisto de una base de datos donde se almacena periódicamente la información relevante.
- Se implementó un sistema de control a través de la comunicación inalámbrica wifi punto a punto de un robot comercial cuya arquitectura está basada en procesadores ATMEL de la familia ATmega. El control se realiza por medio de una plataforma web que puede ser accedida desde un teléfono celular con sistema operativo android como así también desde una PC.

Los resultados esperados para los nuevos temas que se están desarrollando en el proyecto son:

- Se espera para el año en curso obtener los resultados del ahorro energético del sistema y realizar las modificaciones necesarias.

- El objetivo en cuanto al control del sistema robótico, consiste en desarrollar aplicaciones de control inalámbrico de forma remota para proveer soluciones aplicadas a la inspección en ambientes inhóspitos, mapeo y detección de irregularidades.
- Finalmente, se buscará la puesta en funcionamiento del laboratorio remoto diseñado para la materia física 1, en donde se realizará una prueba piloto y se evaluarán las diferentes etapas del sistema.
- En cuanto a la nueva línea, se implementará un sistema dotado de sensores de turbidez, temperatura y conductividad que permita medir en tiempo real la contaminación en arroyos.

4. Formación de Recursos Humanos

Dentro de la temática de la línea de I+D, todos los miembros del proyecto (excepto dos) 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 Doctores en Ingeniería, un Magister y tres docentes realizando la Maestría en temas relacionados.

Una becaria de entrenamiento CIC en temas de eficiencia energética y un becario de vinculación en temas de red de sensores para contaminación en arroyos. Además, se cuenta con una pasante que realizó sus prácticas profesionales en el tema de laboratorio Remoto.

Adicionalmente, se cuenta con la colaboración de estudiantes avanzados.

5. Bibliografía

[1] J. Osio, J. Salvatore , D. Alonso , V. Guarepi , M. Cappelletti , M. Joselevich , M. Morales, “Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”, UNNE, Ciudad de Corrientes, WICC 2018.

[2] 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

[3] Leila Fatmasari Rahman, ”Choosing your IoT Programming Framework: Architectural Aspects”, 2016 IEEE 4th International Conference on Future Internet of Things and Cloud

[4] Osio J., C. Aquarone, E. Hromek, J. Salvatore , “Plataforma de desarrollo para IoT”, IV conaiisi, 2017

[5] Eficiencia energética en los edificios, josé maria Fernandez salgado, AMV ediciones, 2011

[6] Energía Solar Fotovoltaica 2ª Edición, Miguel Aparicio, Edit. marcombo, 2009

[7] Laboratorio remoto en un entorno virtual de enseñanza aprendizaje, M A. Revuelta, FI-UNLP, 2016

[8] 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.

[9] Jan Axelson, “Embedded Ethernet a internet complete”, LLC, 2003

Entorno de Desarrollo Python para EDU-CIAA-NXP

Diego A. Godoy, Eduardo O. Sosa, Juan de Dios Benítez, Hernán Bareiro, Sebastián Paniagua, Luciano Benítez

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

{diegodoy; eduardo.sosa; juan.benitez; hbareiro; spaniagua,lbenitez }@citic.ugd.edu.ar,

Resumen

Este desarrollo se enmarca en 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. Este artículo presenta un estado de situación de un prototipo de IDE multiplataforma, con algunos mecanismos de extensión, que permiten el desarrollo de aplicaciones en lenguaje Python, implementables en la EDU-CIAA.

Palabras claves: IDE; Phyton; OpenSource; EDU-CIAA;

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 (CITIC), aprobado por la Resolución Rectoral 07/A/17.

Entre las líneas de trabajo con entregables producidos se encuentran: a) Construcción de una plataforma de gestión y simulación de datos de redes de sensores inalámbricos, b) interfaz web para el simulador de WSN Shawn, c) sistemas de gestión de residuos de la ciudad de Posadas con tecnologías de Internet de las cosas, d) monitoreo parámetros de proceso en la industria del Té.

Introducción

La CIAA es un proyecto que nace en el año 2013 a través de una iniciativa en conjunto entre la Asociación Civil de Sistemas Embebidos ASCE [1] y la Cámara Argentina de Industrias Electrónicas, Electromecánicas y Luminotécnicas CADIEEL [2] con el objetivo de impulsar el desarrollo tecnológico nacional, ubicar a la electrónica argentina en una posición de liderazgo y generar cambios en la gestión y desarrollo del conocimiento sobre electrónica en el ámbito de las instituciones educativas y las industrias.

Así mismo, las instituciones educativas incluyen en sus planes de estudio contenido relacionado a la programación de computadoras y a la interacción con componentes electrónicos. Conscientes de esta tendencia, las instituciones y empresas que crearon la Computadora Industrial Abierta Argentina (CIAA) [3] diseñaron una versión educativa (EDU-CIAA) "...para lograr un impacto en la enseñanza primaria, secundaria y universitaria." [4]. La EDU-CIAA pretende acercar a alumnos y docentes a la electrónica a través de una plataforma sencilla y es necesario combinarla con un componente de software igualmente sencillo para lograr un producto completo en el contexto de la educación y la enseñanza.

Python [5] es el lenguaje más popular para la enseñanza de cursos introductorios de Ciencias de la Computación [6] debido a su reducida curva de aprendizaje y su capacidad para ejecutarse en múltiples plataformas. En la actualidad se disponible un firmware que permite ejecutar código fuente escrito en lenguaje Python en la EDU-CIAA. El proyecto es llamado micro-python [7] y fue desarrollado para un hardware llamado pyboard. Paralelamente se está desarrollando un editor que permite escribir código fuente en lenguaje Python para desplegarlo en la EDU-CIAA. Si bien existen entornos de desarrollo para micro-python, como uPyIDE [8] y ESCut [9], que cualquier persona pueda acceder y modificar el código fuente, no se tiene previsto construir sus interfaces gráficas de usuario con tecnologías web ni brindar mecanismos de extensión que permitan al usuario adaptar y escalar la herramienta a sus necesidades.

Al considerar lo expuesto en los párrafos anteriores, y teniendo en cuenta la constante evolución del software en torno a la EDU-CIAA, se plantea el diseño e implementación de un prototipo de Entorno Integrado de Desarrollo (IDE) que cuente con una arquitectura abierta a nuevos cambios y/o modificaciones a través de mecanismos de extensión que permita el desarrollo de programas en lenguaje Python para la plataforma EDU-CIAA con las ventajas de una IDE.

Línea de Investigación

En esta línea de investigación se han planteado varios objetivos. El objetivo específico de la misma corresponde a:

Diseñar un prototipo de IDE multiplataforma con mecanismos de extensión para desarrollar aplicaciones en lenguaje Python y desplegarlas en la EDU-CIAA.

Como objetivos específicos se propusieron los siguientes: 1) Establecer el estado actual del Proyecto EDU-CIAA y su relación con los Entornos Integrados de Desarrollo. 2) Analizar las tecnologías que permiten crear aplicaciones multiplataforma que se ejecuten en el escritorio y cuyas interfaces gráficas se desplieguen en navegadores web convencionales. 3) Determinar los requerimientos de software necesarios para diseñar un IDE que permita escribir código fuente en lenguaje Python y desplegarlo en la EDU-CIAA. 4) Diseñar un modelo de dominio de la Interfaz Gráfica de Usuario (GUI) de un IDE utilizando el Lenguaje de Modelado para los Flujos de Interacción [8] (IFML). 5) Diseñar un modelo de dominio de una capa de servicios utilizando el Lenguaje Unificado de Modelado [9] (UML). 6) Desarrollar 2 prototipos de aplicaciones con el lenguaje Python y desplegarlos en la EDU-CIAA utilizando el prototipo de IDE propuesto para luego evaluar el resultado con usuarios finales.

Resultados

En este apartado, se realiza la presentación general del diseño de la solución. Posteriormente, se describirá la arquitectura del sistema sus componentes y cada una de sus capas. En cada capa, se definirán los componentes de hardware y software utilizados, y se establecerán las tareas realizadas y funcionalidad que cada componente. En la Fig. 1 se ilustra el detalle de la implementación, la cual está conformada por una capa de Hardware y una capa de Software.

Fig. 1 Detalles de implementación del IDE

La capa de hardware se compone por una placa EDU-CIAA-NXP que es una versión de bajo costo de la CIAA-NXP pensada para la enseñanza universitaria, terciaria y secundaria (ver Figura 2).

Fig. 2 EDU - CIAA - NXP

La EDU-CIAA está basada en la CIAA-NXP, por ser la primera versión de la CIAA que se encuentra disponible, su microcontrolador es también el LPC4337 (dual core ARM Cortex-M4F y Cortex-M0) como el de la CIAA en la cual se basa. Con el objetivo de abaratar costos y reducir su complejidad la EDU-CIAA incorpora sólo algunas de las funcionalidades de la CIAA. A su vez, con el fin de permitir el desarrollo de algunas prácticas sencillas sin que sea necesario recurrir a hardware adicional, incluye además algunos recursos que no están presentes en la CIAA.,

La capara de software la misma se compone del IDE, dividido en Interfaz de Usuario y la Consola de Servicios, ambas partes del prototipo desarrollado en comunicación con el Sistema Operativo.

La Interfaz de Usuario permite gestionar un proyecto de código fuente Python. Se implementa el diseño resultante de la etapa 2 utilizando las tecnologías web HTML5, CSS y Javascript. El resultado es una aplicación web que, en estrecha colaboración con la Consola de Servicios, permite gestionar un proyecto de código fuente Python.

La consola de servicios: Este componente debe ser diseñado de tal manera que, respetando el estilo arquitectural de REST, abstraiga a la GUI de la complejidad al lidiar con capas subyacentes (sistema operativo, sistema de archivos, etc.). Está implementada con el lenguaje Python y contiene un mecanismo de extensión que permite agregar servicios a la capa. Entre las extensiones que se han implementado encontramos :

- Extensión de lectura de estado de la EDU-CIAA, informando e informar al usuario de forma conveniente.

- Extensión que permite desplegar el proyecto Python en la EDU-CIAA.
- Extensión de gestión de una colección líneas de código frecuentemente utilizadas.

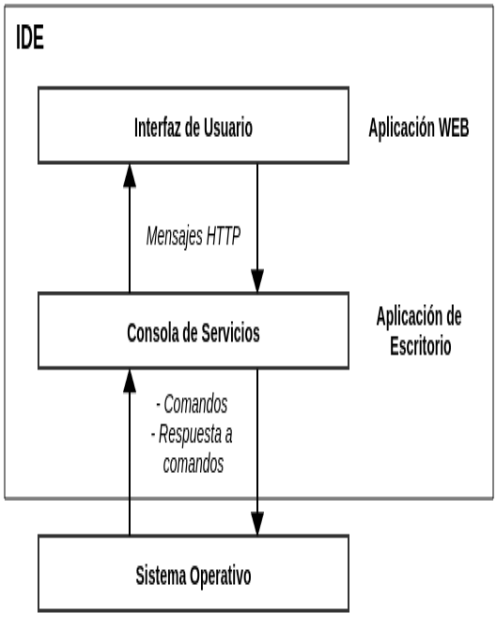


Fig. 3 Detalles de implementación del IDE

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] ASCE. (2015, Aug.) Sitio Oficial. [Online]. http://www.sase.com.ar/asociacion-civil-sistemas-embebidos/

[2] CADIEEL. (2015, Aug.) Sitio Oficial. [Online]. http://www.cadieel.org.ar/esp/index.php

[3] CIAA. (2015, Aug.) Sitio oficial. [Online]. http://www.proyecto-ciaa.com.ar/

[4] EDU-CIAA. (2018, Aug.) Documentación. [Online]. http://www.proyecto-ciaa.com.ar/devwiki/doku.php?id=desarrollo:edu-ciaa:edu-ciaa-nxp

[5] Python. (2015, Aug.) Sitio oficial. [Online]. https://www.python.org/

[6] Philip Guo. (2015, Aug.) Python is Now the Most Popular Introductory Teaching Language at Top U.S. Universities. [Online]. http://cacm.acm.org/blogs/blog-cacm/176450-python-is-now-the-most-popular-introductory-teaching-language-at-top-us-universities/fulltext

[7] The MicroPython forum. (218) MicroPhyton. [Online].

https://micropython.org/

[8] Piero Fraternali Marco Brambilla, *Interaction Flow Modeling Language: Model-Driven UI Engineering of Web and Mobile Apps with IFML*.: Morgan Kaufmann Publishers, 2015.

[9] UML. (2018, Aug.) [Online]. http://www.uml.org/

[10] CIAA. (2018, Aug.) Sitio oficial. [Online]. http://www.proyecto-ciaa.com.ar/

[11] Python. (2018, Aug.) Sitio oficial. [Online]. https://www.python.org/

Laboratorios con simulación y con equipo real en la enseñanza de redes de computadoras en el nivel universitario

Daniel Arias Figueroa, Ernesto Sánchez, Gustavo Gil, Loraine Gimson, Álvaro Gamarra, Rodolfo Baspineiro, Romina Chacón, Agustín Colque

C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada)
Facultad de Ciencias Exactas de la Universidad Nacional de Salta
Facultad de Ingeniería de la Universidad Católica de Salta
Av. Bolivia Nº 5150 – Campo Castañares – Salta – Tel. 0387-4255476
daaf@cidia.unsa.edu.ar, gdgil@unsa.edu.ar

Resumen

El presente trabajo es una continuación del proyecto acreditado que finalizó en diciembre del 2018 "Estudio de la influencia del uso de simulación en la enseñanza de redes de computadoras en el nivel universitario", en el cual se contrasto la enseñanza tradicional de redes con el uso de simulación. Allí se evidenció, que los estudiantes que utilizaron simulación mostraron un nivel de comprensión del concepto evaluado más alto que aquellos que siguieron un procedimiento tradicional, además de disponer de una buena actitud hacia el estudio de la asignatura.

Nuestra línea de investigación, pretende evaluar la influencia de la utilización de software de simulación, en la enseñanza de contenidos de redes de computadoras en carreras de grado y cursos de postgrado a fin de contrastar con la enseñanza con laboratorios con equipo real Cisco y Mikrotik. Para ello nos planteamos un estudio aplicado, longitudinal y experimental que, considera aspectos cuantitativos y cualitativos, que permitan determinar si los estudiantes que realizan prácticas en un entorno de simulación aprenden significativamente más que aquellos estudiantes que se capacitan con laboratorios de equipos reales de red.

Palabras clave: protocolo tc-pip, emulación, simulación, enseñanza, redes.

Contexto

La línea de investigación se encuentra apoyada por el C.I.D.I.A. (Centro de Investigación y Desarrollo de Informática Aplicada) que depende de la Facultad de Ciencias Exactas de la Universidad Nacional de Salta y por la Facultad de Ingeniería de la Universidad Católica de Salta, por lo tanto se cuenta con toda la infraestructura disponible para esta investigación. El proyecto contará con el financiamiento del CIUNSa – Consejo de Investigación de la Universidad Nacional de Salta y el Consejo de Investigación de la Universidad Católica de Salta.

Introducción

Los conceptos y fundamentos de redes son difíciles de asimilar debido a la complejidad de los procesos involucrados que no son siempre visibles. Esto, sumado a los costos elevados de equipos específicos necesarios para montar un laboratorio de red, y a los escasos recursos con los que cuentan las Universidad Nacionales, 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 clase (Laboratorio virtual).

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, sino 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.

Así, las computadoras personales, con la variedad de software que se ha desarrollado, tienen en sí mismas un gran potencial para mejorar el proceso de enseñanza-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.

Líneas de Investigación, Desarrollo e Innovación

Esta investigación pretende dar respuesta a la pregunta, ¿cuál es la influencia del uso de herramientas de simulación en el aprendizaje de conceptos y fundamentos sobre redes de computadoras en estudiantes del nivel universitario? De esta pregunta principal se derivan las siguientes preguntas:

- ¿Existe diferencia estadísticamente significativa en el nivel de comprensión de los conceptos que se van a a evaluar sobre redes IP, entre estudiantes que reciben instrucción mediada por herramientas de simulación y estudiantes que reciben instrucción utilizando laboratorios con equipos reales de red?

- ¿Existe relación entre la utilización de herramientas de simulación y la actitud de los estudiantes hacia el aprendizaje de las asignaturas en cuestión y de otras asignaturas del plan de estudio de la carrera Licenciatura en Análisis de Sistemas?

Es importante destacar que, en el marco del proyecto de investigación, se realizarán experiencias con temáticas tales como: enrutamiento IP estático, enrutamiento IP dinámico con el protocolo RIP, protocolo DNS, protocolo ARP, Direccionamiento IP y otros. Todas estas temáticas corresponden a las asignaturas Redes de Computadoras I y Redes de Computadoras II del plan de estudios de la carrera Licenciatura en Análisis de Sistemas de la UNSa.

Para ello se llevará a cabo un estudio aplicado, longitudinal y experimental que, asumiendo la complejidad del fenómeno educativo, considerará aspectos cuantitativos y cualitativos, contrastando la enseñanza con simulación con la enseñanza con equipo real de redes CISCO y MIKROTIK.

Asimismo, la investigación persigue en el marco del paradigma cualitativo los siguientes objetivos:

- Realizar entrevistas y estudios de casos con los alumnos durante los períodos previstos.
- Diseñar actividades prácticas y de investigación adecuadas para ser realizadas con los simuladores y con laboratorio con equipo real.
- Fomentar el aprendizaje cooperativo entre los estudiantes.
- Introducir a los estudiantes de licenciatura en la cultura científica actual, que concibe la simulación por computadora como una herramienta fundamental para el estudio, la investigación y la experimentación.
- Contratar el resultado producto de la investigación con estudios similares en otras universidades.

Resultados y Objetivos

El objetivo general del estudio aquí planteado es determinar el impacto del uso de herramientas de simulación en el aprendizaje de conceptos y fundamentos de redes de computadoras en estudiantes del nivel universitario. A partir de este objetivo se plantearon los siguientes objetivos específicos:

- Establecer la relación entre la utilización de la simulación y la actitud de los estudiantes hacia el aprendizaje de las redes de computadoras y de temáticas de otras asignaturas del plan de estudios;
- Determinar si existe diferencia significativa en el nivel de comprensión de los principios de redes de computadoras, entre estudiantes que reciben instrucción mediada por herramientas de simulación y estudiantes que reciben instrucción con laboratorios con equipo real.

En esta primera etapa del proyecto se realizó el relevamiento, análisis y selección de las herramientas de simulación que se utilizarán en la experimentación. También se están desarrollando las guías prácticas para utilizarse con el emulador de red GNS3. Se pudo realizar una experiencia piloto que nos permitió una primera validación del material elaborado.

Esta planificado realizar durante el año 2019 las primeras experiencias en el aula en el marco de las asignaturas Redes de Computadoras I y Redes de Computadoras II de la Lic. en Análisis de Sistemas y Redes I y Redes II de la Ingeniería en Telecomunicaciones de la Universidad Católica de Salta.

Formación de Recursos Humanos

El equipo de investigación se conforma con un director un co-director y seis miembros entre docentes, egresados y estudiantes avanzados. Además se sumaran docentes de otras Universidades del país a través de convenios de cooperación que se firmarán oportunamente.

- Especialidad en Seguridad y Redes – UNLP. “Virtualización para la enseñanza de redes IP”. Miguel A. Aguirre. (En desarrollo).
- Cursando la Maestría en Redes de Datos en la UNLP. Ing. Álvaro I. Gamarra.
- Cursando la Licenciatura en Análisis de Sistemas en la UNSa. Romina Chacón y Agustín Colque.

Referencias

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.

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: Univ. de Salta.*

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.

CÁMARA, ALZUGARAY (2011). Trabajos Prácticos, Métodos de Simulación y aprendizaje significativo. La Tecnología Educativa al servicio de la Educación Tecnológica.

ARIAS FIGUEROA, D. SÁNCHEZ, E. Colaboración de BASPINEIRO (2018). Redes de Computadoras II con Packet Tracer. Editorial de la Universidad Nacional de Salta – EUNSa – ISBN 978-987-633-527-0; 1a ed. - Salta - E-Book - CDD 004.678.

ARIAS FIGUEROA, D. (2015). Redes de Computadoras I con Packet Tracer. Editorial de la Universidad Nacional de Salta – EUNSa – ISBN 978-987-633-132-6; 1a ed. - Salta - E-Book - CDD 004.68.

Selección de Sistemas Operativos apropiados para extender la vida útil de residuos electrónicos (RAEE)

Eduardo Rodríguez[1], Claudia Deco[1,2], Luciana Burzacca[1,2], Santiago Costa[1], Cristina Bender[1,2], Cristian Del Giglio[1]

Facultad de Química e Ingeniería del Rosario, Universidad Católica Argentina
Universidad Nacional de Rosario
{ejrodriguez, cdeco, lburzacca, santiagocosta, cbender}@uca.edu.ar

Resumen

En la actualidad los gobiernos, empresas e individuos se encuentran ante la problemática de los residuos electrónicos (RAEE). Anualmente se generan millones de desperdicios dentro de esta categoría y la velocidad con la que avanza la tecnología hace que diariamente se incorporen mayor cantidad de dispositivos. Existen oportunidades para que dispositivos desechados, luego de un tratamiento adecuado, puedan ser reutilizados con distintos fines. Dentro de este proyecto se propone seleccionar, analizar y testear distribuciones de sistemas operativos que extiendan la vida útil de los equipos de computación personal permitiendo así su reutilización con el fin de lograr un impacto positivo en la sociedad y al mismo tiempo reducir los RAEE. Los resultados obtenidos hasta el momento, muestran que existen distribuciones de sistemas operativos adecuadas para enfrentar el desafío planteado en computadoras de más de ocho años de antigüedad.

Palabras Claves: RAEE, Sistemas Operativos, Reutilización de PCs.

Contexto

Esta línea de I+D se está llevando a cabo a través de un proyecto del Departamento de Investigación Institucional de la Facultad de Química e Ingeniería del Rosario de la Universidad Católica Argentina. El proyecto involucrado es: PID UCA "Propuesta de un Sistema de Gestión Integral de Residuos Electrónicos" (2017 - 2020).

Introducción

En la Argentina todavía existe poca conciencia sobre la importancia que posee el tema de los Residuos Electrónicos (RAEE). Este tipo de residuos tiene un crecimiento exponencial y constante. Produce un gravísimo impacto ambiental y pérdida económica ocasionada a partir de su no reutilización y su no reciclado. A los fines de poder proponer soluciones al problema se estudian el impacto ambiental que implica no tratar el residuo, la elaboración de propuestas de legislación específica, el diseño de sistemas integrales de gestión, como también el estudio de proyectos de emprendimientos logísticos y de tratamiento para la valorización de los mismos, que resulten sustentables y viables económicamente, e inspirados en las mejores prácticas de gestión ambiental existentes a nivel mundial.

El promedio mundial de RAEE generado por habitante por año se estima en 7 kg, estando actualmente la Argentina por sobre el promedio en unos 8,5 kg y se estima en 15 kg para el año 2025 [1]. Muchos de los equipos descartados están

en funcionamiento, y pueden tener una extensión de su vida útil si son recuperados a tiempo.

Los RAEE constituyen una verdadera Mina de Oro Tóxica que combina un alto valor de algunos de sus componentes con la peligrosidad de otros. El contenido de las plaquetas electrónicas y procesadores son un 300% más rico en cobre y otros metales preciosos que el promedio de las mejores vetas encontradas en la minería tradicional, lo que ha inaugurado el concepto de Minería Urbana. También se recuperan plásticos, vidrio y metales ferrosos [2]. Por otro lado la presencia de elementos como mercurio, cadmio, plomo, bromo, selenio, bifenilos policlorados, policloruros de vinilo e ignífugos como el arsénico y el amianto, hace de ellos de gran poder contaminante llegando a aportar el 70% de metales pesados a los rellenos sanitarios. La mayor contaminación de los RAEE se produce cuando se queman o mojan, produciendo el lixiviado hacia las napas subterráneas, contaminando tierra, aire y agua [2]. Por eso, el símbolo que los representa internacionalmente es enfático: Los RAEE no deben ir a la basura.

Se estima que el 50% de estos residuos están arrumbados en oficinas, hogares, entes públicos o depósitos; más del 40% se entierra o se descarta en basurales y rellenos y cerca del 10% ingresa en esquemas informales o formales de gestión de residuos [3, 4].

En este proyecto se propone el estudio de esta actividad en el contexto de un grupo de investigaciones que permita:

- La colaboración interinstitucional e interdisciplinaria entre especialistas de cada campo.
- Conocer la situación actual de RAEE en la región.
- Estudiar e investigar soluciones al problema del RAEE a nivel mundial aplicables a nuestra realidad regional.
- Analizar y testear distribuciones de Sistemas Operativos que extiendan la vida útil de equipos permitiendo su reutilización.
- Proponer un Sistema Integral de Gestión de RAEE.

La realización de este proyecto permite traducir el conocimiento generado en investigación y en un importante desarrollo tecnológico para un sector de alta relevancia como es recuperar equipos y realizar una reinserción responsable de los mismos en la sociedad.

Líneas de Investigación, Desarrollo e Innovación

Actualmente, en la región de la provincia de Santa Fe, con una población de 2.024.478 y según el último censo realizado por el INDEC en 2010, sólo un 46,31% tiene acceso a equipos de computación en su hogar, lo que representa 473.796 hogares. Si llevamos estos números a nivel país, los mismos se encuentran en línea con la provincia con sólo 46,86% de hogares que poseen computadoras. En caso de citar a habitantes de todo el país que dicen utilizar computadoras frecuentemente se observa un ligero crecimiento llegando al 52,82%. Esto demuestra la enorme oportunidad que existe en el campo de la reutilización de equipos informáticos que, con un tratamiento, gestión y logística adecuadas, sumadas a la correcta selección de un Sistema Operativo que maximice y extienda su vida útil, podrían reinsertarse en distintos ámbitos y tener un impacto positivo en la sociedad.

La velocidad con la que avanzan las tecnologías informáticas hace que los componentes tecnológicos tengan un alto grado de obsolescencia tecnológica en un corto período. Actualmente se estima que las empresas, para mantenerse

competitivas y actualizadas, con un entorno seguro y con usuarios con un nivel adecuado de productividad, reemplazan sus equipos de computación personal en un plazo de entre tres y cuatro años [5, 6].

Según estudios de Camoca [7] se sabe que en la actualidad los residuos electrónicos en Argentina ascienden a 21.010.000 de equipos relacionados a computación personal. Si tomamos como punto de partida el año 2015 se ve un crecimiento de 10% en equipos en desuso en 2016 y un 13% en 2017, esto muestra un incremento anual en la cantidad de equipos en desuso. Si realizamos el análisis en kilos de material los porcentajes también son crecientes año tras año, 5% para los años 2015 al 2016 y 11% entre 2016 y 2017.

El desafío es, entonces, darle una disposición final a los equipos de acuerdo a los marcos legales que actualmente se encuentran en vigencia.

Un punto importante a considerar es que las empresas dedicadas al tratamiento de residuos además de la capacidad instalada tienen otra limitante, que es que son aranceladas, por lo cual las empresas y gobiernos se ven en la obligación de afrontar un costo cuando podrían avanzar con donación de equipos si dotaran a los mismos de un Sistema Operativo capaz de recuperar el potencial de los equipos que en otros ambientes han llegado a un nivel de obsolescencia determinado.

En este proyecto se pretende analizar distintas alternativas de sistemas operativos que permitirán recuperar equipos de la obsolescencia y realizar un ranking que destaque y clasifique los mismos. Asimismo, se propone realizar tests de performance que demuestren las distintas características de los SO seleccionados. Para esto, se preparó un laboratorio de computadoras con distintas arquitecturas y componentes de hardware.

Hasta el momento, se analizaron 28 distribuciones de Linux entre las cuales se seleccionaron: Bodhi Linux 5.0.0-64, Puppy XenialPup 7.5, Elive 3.0.2 y Mint 19 "Tara" - Xfce (32-bit). Estos SO fueron seleccionados teniendo en cuenta características esenciales para el desarrollo del trabajo y la recuperación de hardware obsoleto. Entre éstas se encuentran:

- Soporte para 32 bits, una determinada gama de equipos sólo soporta este tipo de arquitectura.
- Diversidad de versiones, en todos los casos las distribuciones se pueden ejecutar en modo full, es decir instalado completamente en el disco rígido del computador como también está disponible para su ejecución mediante un medio físico como CD, DVD o USB, que también incluyen soporte para otras arquitecturas como x64 y algunas versiones pasadas.
- Requerimientos de hardware, conforme el objetivo que se desea probar, se seleccionan distribuciones de SO que requieran mínimas especificaciones y permitan operar de manera óptima y estable.
- Versiones actualizadas, contar con el apoyo constante de las comunidades y un producto actualizado, es un aspecto de vital importancia por diversos motivos, como ser, mantener el entorno seguro ante vulnerabilidades y ataques. Acceder a mejoras del producto relacionadas, pero no limitadas a interfaz, nuevos repositorios de software, mayor cantidad de drivers, estabilidad y robustez del sistema basado en la experiencia de la distribución en mercado y cantidad de usuarios.
- Documentación, otro aspecto priorizado, dado que contar con manuales adecuados, guías de trabajo, instructivos y foros instrumentados

como un lugar para poder afrontar problemas, es un diferencial muy valioso.

Con los cuatro sistemas operativos seleccionados para evaluar, se desarrollaron pruebas focalizadas en las siguientes categorías, considerando los aspectos indicados en cada categoría:

- Experiencia del usuario: Simplicidad para la descarga, Documentación de soporte disponible, Facilidad de instalación, Interfaz de usuario, Asistente para configuración de funcionalidades, y Variedad y Calidad de aplicaciones incluidas en la distribución.
- Estabilidad: Cantidad de operaciones por hora sin reinicio, Cloud computing y sincronización con Google Drive & Sync, Disponibilidad de versiones actualizadas, Disponibilidad de seguridad incluida (firewall, antivirus, otros), Soporte para UEFI (Unified Extensible Firmware Interface).
- Performance: Cantidad de versiones disponibles, Utilización del CPU, Utilización de memoria, Utilización de disco (I/O), Utilización de red, y Navegación por Internet.

Resultados y Objetivos

El objetivo general de este proyecto es la elaboración de propuestas de acción para la puesta en marcha de un Sistema Integral de Gestión de RAEE para la región. Para esto se tienen los siguientes objetivos específicos:

- Establecimiento de línea de base e indicadores de gestión.
- Propuesta de marco legal determinando las responsabilidades de los actores involucrados.
- Elaboración de Proyectos de Unidades de Gestión de RAEE para reutilizado y reciclado.
- Elaboración de una propuesta de recolección, logística y acopio de RAEE.
- Análisis económico-financiero del sistema integral.
- Modelización y optimización del sistema general.
- Análisis de SO orientados a la reutilización de equipos.

En la actualidad existe una gran cantidad de equipos informáticos dispuestos por grandes empresas y gobierno, que luego de un extensivo período de uso no se encuentran en condiciones de ser reinsertados en ambientes productivos, dado que pasado su ciclo de vida útil la performance de estos se torna prácticamente nula. Esto genera el análisis y búsqueda de SOs capaces de maximizar los recursos disponibles en equipos de computación personal de hasta catorce años de antigüedad.

La evaluación de los resultados obtenidos se está realizando mediante un análisis cuantitativo y cualitativo y como parte de la solución se está analizando el mercado actual de SOs pre-seleccionando aquellos con características enfocadas en un mínimo consumo de recursos y un máximo aprovechamiento de las capacidades de hardware con más de diez años de antigüedad. En particular, del análisis de los distintos SO se seleccionaron cuatro distribuciones: Puppy XenialPup 7.5, Elive 3.0.2, Bodhi 5.0 apppack, Linux Mint 19 Tara - Xfce (32-bit). De acuerdo a la experimentación realizada se determinó a Linux Mint 19 Tara - Xfce (32-bit) como la más óptima. Así mismo, es válido destacar que las distribuciones Puppy XenialPup 7.5 y Elive 3.0.2 también cumplirían con lo planteado en este proyecto, aunque con la limitante de acotar su uso a hardware de no más de diez años de antigüedad para mantener una óptima performance.

Formación de Recursos Humanos

El equipo de trabajo está integrado por los Magisters Eduardo Rodríguez y Cristina Bender, la Doctora Claudia Deco, la especialista Luciana Burzacca, y los Licenciados Santiago Costa y Cristian Del Giglio, todos investigadores de la Universidad Católica Argentina.

En particular, el Lic. Costa está cursando la Maestría en Redes de Datos en la Universidad de La Plata y el Lic. Del Giglio logró su título en el año 2018 con la tesina "Estudio y selección de sistemas operativos low resources".

Además, los conocimientos adquiridos se vuelcan en la comunidad académica a través de las actividades docentes que los investigadores de este proyecto realizan en las carreras Ingeniería Industrial, Ingeniería Ambiental y Licenciatura en Sistemas. Asimismo, es intención incorporar alumnos de estas carreras y proponer subproblemas identificados durante el trabajo de investigación de este proyecto para la realización de tesinas y proyectos finales de grado.

La conformación del grupo de investigación por docentes de la carrera de Ingeniería Industrial, la carrera de Ingeniería Ambiental y de la carrera de Licenciatura en Sistemas permite un enfoque multidisciplinario y adecuado de esta problemática. La viabilidad del proyecto está sustentada en los avances obtenidos mediante trabajos previos de los integrantes en el área y su interacción con otros grupos de investigadores tanto latinoamericanos como europeos.

Referencias

[1] Fernández Protomastro, G. (2014). Buenas Prácticas para la gestión de los residuos de aparatos eléctricos y electrónicos - RAEE. 1a ed. Ciudad Autónoma de Buenos Aires: Grupo Uno. 178 p. ISBN 978-987-29862-3-0

[2] Fernández Protomastro, G. (2013). Minería Urbana y la Gestión de los Residuos Electrónicos. - 1a ed. Ciudad Autónoma de Buenos Aires: Grupo Uno. 317 p. ISBN 978-987-29862-1-6

[3] Greenpeace, (2011). Basura Electrónica, la otra cara de la tecnología. Disponible en: http://www.greenpeace.org/argentina/Global/argentina/report/2011/contaminacion/basura_electronica_otra_cara_tecnologia.pdf. Consultado 07/04/2018.

[4] Greenpeace, (2012). Minería y Basura Electrónica, el manejo irracional de los recursos. Disponible en: http://www.greenpeace.org/argentina/Global/argentina/report/2012/contaminacion/inform-raee-V-1.pdf Consultado 07/04/2018.

[5] [(2) National Intruments, "Políticas de Ciclo de Vida de Productos de Hardware", http://www.ni.com/life-cycle/esa/hardware.htm 3/12/2018,

[6] (3) Rowan Gonzalez, "The Computer Life Cycle", 25/2/2013, https://www.lifeder.com/ciclo-vida-computadora 3/12/2018].

[7] Camoca. Disponible en: http://www.camoca.com.ar/

[8] Ecotic (2017) Disponible en: http://www.ecotic.es/es/246571/Documentos.htm consultado 07/04/2018.

[9] Magalini, F.; Kuehr, R.; Peter Baldé, C. (2015) eWaste en América Latina. Análisis estadístico y recomendaciones de política pública. GSMA. Instituto de la ONU para el Estudio Avanzado de la Sostenibilidad. Disponible en: http://www.gsma.com/latinamerica/wp-content/uploads/2015/11/gsma-unu-ewaste2015-spa.pdf. Consultado 07/04/2018.

Site survey para la implementación de una red WI-FI

César Alejandro Castillo, José Federico Medrano, Mario Alberto Tejerina, González Soledad Carolina
Facultad de Ingeniería, Universidad Nacional de Jujuy
Contacto: ce.al.castillo@gmail.com

RESUMEN

En este trabajo se propone una alternativa para la realización de estudios de sitio a bajo costo, para la implementación de una red Wifi en las instalaciones de la Facultad de Ingeniería, analizando los parámetros más importantes para determinar la colocación apropiada de los dispositivos inalámbricos con el objetivo de obtener la red deseada, evitando interferencias, reduciendo zonas de cobertura conflictivas, evitando que los dispositivos se desconecten y perjudiquen a la red completa.

CONTEXTO

La presente investigación corresponde a una Práctica Profesional Supervisada de fin de carrera, realizada por Alumnos avanzados de la Facultad de Ingeniería de la Universidad Nacional de Jujuy de las carreras de Licenciatura en Sistemas e Ingeniería en Informática. El equipo está conformado por Alumnos y Profesores de dicha institución.

1. INTRODUCCIÓN

La prospección del sitio o Site Survey nos permite ubicar estratégicamente los puntos de acceso en un lugar determinado y principalmente para asignar los canales de frecuencia correctamente, evitando de esta manera la interferencia entre dispositivos. Además hay que tener en cuenta que si un dispositivo se conecta en un punto de acceso desde una distancia en la cual la potencia de la señal es baja, este se desconecta, por lo tanto la solicitud de conexión y reconexión ralentiza toda la red perjudicando a todos los usuarios conectados a ese punto de acceso [2]. Entonces es necesario limitar la celda de cobertura hasta un cierto nivel, no permitiendo a los dispositivos conectarse cuando la potencia sea demasiada baja forzandolos a que busque otra área de cobertura sin afectar a toda la red. Es por esto que la planificación de estos tipos de redes es muy importante, porque permite analizar una serie de parámetros que de acuerdo a los requerimientos del cliente, se deberán configurar correctamente para lograr cumplir con sus expectativas [5].

Es posible combinar herramientas para obtener un Site Survey completo sin la necesidad de comprar dispositivos costosos, tales como GPS o medidores de espectros electromagnéticos. Para esto existen dos herramientas, cuya principal ventaja es que son libres y por lo tanto es posible realizar un Site Survey con un alto grado de efectividad a muy bajo costo, estas herramientas son Radio Mobile y Ekahau HeatMapper. Con ellas se pueden utilizar mapas y a partir de estos realizar el estudio de las coberturas.

2. LINEAS DE INVESTIGACION Y DESARROLLO

“Un sondeo del sitio es un proceso tarea a tarea por el cual el sondeador estudia la instalación para conocer el comportamiento, descubre las áreas de cobertura, comprueba si hay interferencias y determina la colocación

apropiada de los dispositivos inalámbricos". [6]. En una red inalámbrica, pueden surgir muchos problemas que impidan que la señal de radio frecuencia llegue a todas las partes de la instalación, tales como interferencias, retransmisiones, latencia, pérdida de paquetes, pérdida de velocidad, menos clientes conectados. Para solucionar estos problemas, se necesita encontrar las regiones donde ocurren. Un sondeo del sitio ayuda a definir los contornos de la cobertura en una instalación concreta y descubrir las regiones donde puede ocurrir una distorsión de trayectorias múltiples y las áreas donde las interferencias son altas, y a buscar soluciones para eliminar tales problemas. Un sondeo del sitio que determina el área de cobertura en una instalación también ayuda a elegir el número de dispositivos inalámbricos que una empresa necesita para cumplir sus requisitos comerciales.

Herramientas Útiles

La complejidad del estudio del Site Survey dependerá del inmueble, de los requerimientos del usuario y de las aplicaciones, cuando un sitio requiere varios Access Points el Site Survey es recomendado. A continuación se estudiará una herramienta útil que nos ayudará a realizar este estudio de forma más sencilla:

Ekahau HeatMapper: Es una herramienta de software libre, para el mapeo de la cobertura rápida y fácil de la tecnología Wi-Fi (802.11). Permite la localización de todos los puntos de acceso. Esta herramienta ayudará a dibujar un mapa de intensidad y posición de las conexiones usando un plano existente o una cuadricula vacía. Para realizar la localización de los puntos de acceso de un entorno hay que empezar a caminar por él, con un computador portátil. [4].

Radio Mobile: Es un programa de simulación de radio propagación que utiliza datos de elevación del terreno que se descargan gratuitamente de Internet para crear mapas virtuales del área de interés, vistas estereoscópicas, vistas en 3-D y animaciones de vuelo. [1]. Usa el modelo digital de elevación del terreno para el cálculo de la cobertura, e indica la intensidad de la señal recibida en varios puntos a lo largo del trayecto.

3. RESULTADOS EXPERIMENTALES

A continuación se realizarán algunas pruebas experimentales a las herramientas descriptas anteriormente, comprobando la efectividad de los servicios que nos proporcionan para realizar un Site Survey y determinar las situaciones posibles en las cuales serán de mayor utilidad.

Se realizó la medición de tres puntos de acceso, ubicándose en un punto en la grilla de la herramienta Ekahau HeatMapper, tomando como referencia la ubicación actual y como resultado se obtiene la cobertura que tiene cada punto de acceso. Para obtener las mediciones de las potencias de los transmisores (dBm), también se puede caminar llevando la notebook con la red Wi-Fi habilitada y no necesariamente conectada a alguna red, ya que Ekahau HeatMapper captura la cobertura de todos los puntos de acceso que logra identificar. Se puede trabajar sobre la grilla que nos proporciona por defecto la herramienta o sobre un mapa o plano del área de la cual se desea realizar el Site Survey. En la siguiente figura se puede observar las ubicaciones de los tres puntos de acceso y la ubicación del usuario.

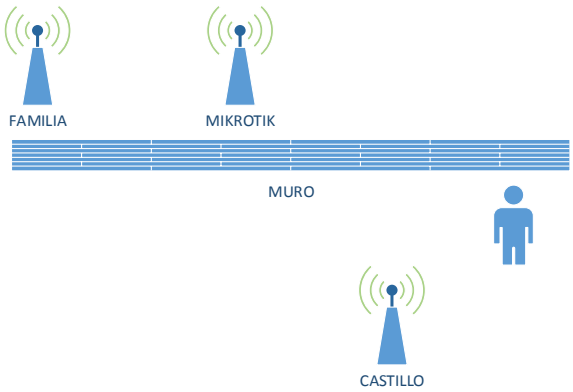


Figura 1. Ubicación de los puntos de acceso a medir

Tener en cuenta la ubicación del muro, el cual influye en la dispersión de las ondas electromagnéticas, generando cierto nivel de atenuación de la señal. En las siguientes figuras se observan las mediciones de las potencias de los transmisores (dBm).

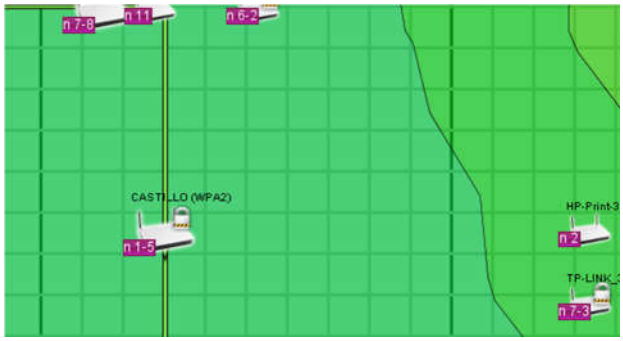


Figura 2. La red con el SSID CASTILLO -35 o mayor dBm

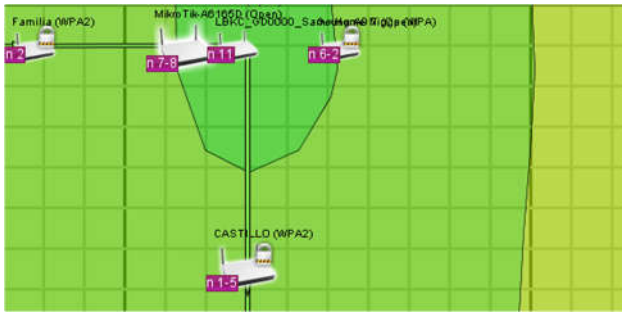


Figura 3. La red con el SSID Mikrotik -48 - -40 dBm

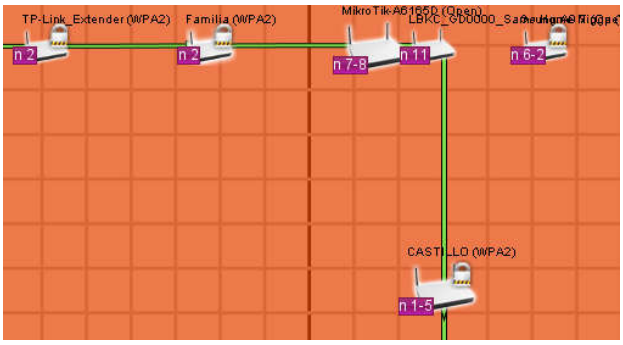


Figura 4. La red con el SSID Familia -88 - -80 dBm

En las mediciones anteriores podemos determinar la cobertura de cada punto de acceso con respecto al punto de ubicación del usuario, en la red con SSID CASTILLO de la figura 2 se puede observar que tiene una cobertura óptima (color verde helecho), mientras que la cobertura del punto de acceso con SSID Mikrotik es menor, y en la red con SSID Familia la cobertura es casi nula. Al alejarse el usuario, la cobertura con respecto al SSID CASTILLO va disminuyendo como se observa en la figura 3 hasta ser nula (figura 4).

A partir de esto también se puede determinar el nivel de interferencias y configurar la selección de canales, los cuales cada punto de acceso debe tener uno diferente [3]. Para el uso de esta herramienta es mejor utilizar un mapa o plano del lugar para identificar las áreas donde ubicaremos los puntos de acceso.

Una herramienta útil para representar un esquema e identificar las antenas o puntos de acceso es Radio Mobile, la cual nos permitirá a través del uso de mapas representar la ubicación de antenas de un área determinada. Estos mapas se pueden descargar de internet con datos de elevaciones del proyecto de la NASA Shuttle Terrain Radar Mapping Misión (SRTM), que provee datos de altura con una precisión de tres segundos de arco (100m). Para descargar un

mapa es necesario ubicar la ciudad o un lugar específico señalando la latitud y longitud, las cuales podemos utilizar la herramienta Google Earh, para encontrar la ubicación exacta.

Radio Mobile nos permite combinar sus mapas con otros mapas topográficos, mapas de carreteras o imágenes satélite, como por ejemplo, mapas de Google Earth. En la figura siguiente se observa la búsqueda de los parámetros de latitud y longitud en un mapa de Google Maps, por ejemplo la ubicación de la Facultad de Ingeniería de la Universidad Nacional de Jujuy en la ciudad de San Salvador de Jujuy. Para ellos capturamos los parámetros Latitud y Longitud haciendo un clic sobre un punto determinado.

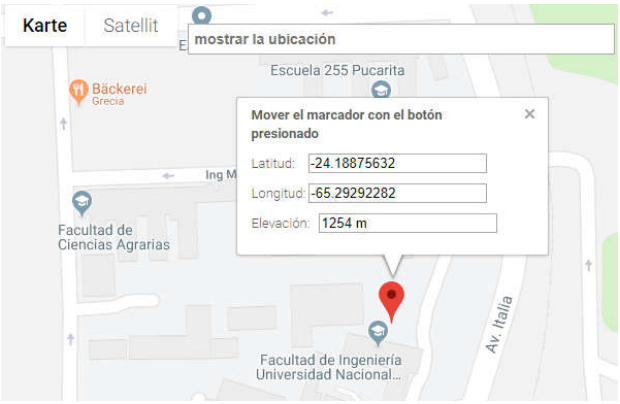


Figura 5. Búsqueda de Latitud y Longitud en Google Maps

Con estos parámetros solicitamos a Radio Mobile que descargue los mapas necesarios, le indicamos el radio que se utilizará y se combinará con otro mapa como puede ser, por ejemplo, uno de Google Earth para tener mejor ubicación a través de otros parámetros tales como calles, carreteras, etc.

Luego se debe crear las redes indicando a la aplicación, las antenas o puntos de acceso que se necesitan y en que parte se ubicarán, indicando para esto la longitud y latitud, parámetros relacionados al tipo de dispositivo que ubicaremos en ese punto y hasta la selección de un icono para representar la antena u otro dispositivo.

En la siguiente figura se realizó la representación de un sondeo ficticio en el predio de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, en el cual se indica la ubicación de tres puntos de acceso que pertenecen a una misma red. Para ellos le indicamos a Radio Mobile los parámetros de latitud, longitud, el tamaño del mapa (en kilómetros) y el tamaño del mapa en pixeles.

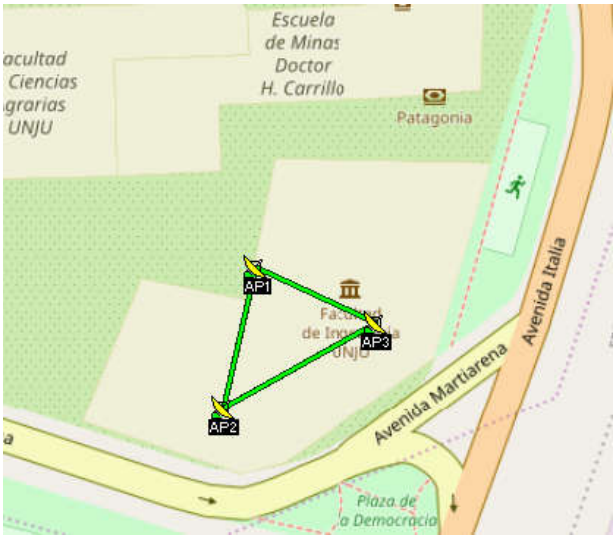


Figura 6. Ubicación de tres puntos de acceso en el predio de la Facultad de Ingeniería

Esta herramienta presenta una grilla que nos indica la altitud en metros de la ubicación en la cual se indique sobre el mapa, estos datos son muy útiles en el caso que debamos emplazar una antena o torres para las cuales es necesario contar con los datos de elevación.

Figura 7. Grilla que indica los datos de elevación.

Estos resultados son experimentos a modo de iniciar la investigación, pero se tiene como objetivo realizar un Site Survey a las instalaciones de la Facultad de Ingeniería de la Universidad Nacional de Jujuy para su posterior implementación de una red Wi-Fi.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto forma parte de la Secretaría de Extensión, Coordinación y Planificación de la Universidad Nacional de Jujuy, la cual propone actividades de Prácticas Profesionales Supervisadas para carreras vinculadas a la informática. Estas prácticas se llevan a cabo mediante la conducción técnica de profesionales informáticos de dicha institución y la participación de tres estudiantes avanzados de la carrera Ingeniería en Informática y/o Licenciatura en Sistemas.

5. BIBLIOGRAFÍA

[1] Grupo de Radiocomunicación, Departamento SSR. “Tutorial de Radio Mobile”, Febrero de 2007.

[2] Paz, J. M., Redes Unificadas de alta densidad CAPsMAN, MUM 2017, Ecuador, 2017.

[3] Peña Millahual, C. A., "Redes Wi-Fi en entornos Windows. Guía práctica de aprendizaje", Revista User, ISBN 978-987-1857-64-7. 2012.

[4] Vanhatupa, T., " Wi-Fi Capacity Analysis for 802.11ac and 802.11n: Theory & Practique", Ekahau. https://www.ekahau.com/es, 2015.

[5] WNDW, "Redes inalámbricas en países en desarrollo", http://wndw.net, ebook, 2013.

[6] CISCO, “Preguntas más Frecuentes sobre Wireless Site Survey”, https://www.cisco.com/c/es_mx/support/docs/wireless-mobility/wireless-lan-wlan/68666-wireless-site-survey-faq.html#qa1, Consultado en Agosto de 2013.

Técnicas para el despliegue de IPv6 en redes LAN

Germán Eduardo Jerez, Víctor José López, Víctor Manuel Longo, María del Pilar Gálvez, Marcelo Alejandro Gallardo, Raúl Alfredo Zamora

Universidad Católica de Santiago del Estero – Departamento Académico San Salvador
Lavalle 333 San Salvador de Jujuy – 0388 4236139
germanjerez@yahoo.com.ar

RESUMEN

El crecimiento de Internet registrado a lo largo de los, aproximadamente, últimos 20 años trajo consigo el problema del agotamiento de las direcciones IPv4 (Internet Protocol versión 4). Como solución a este problema, el IETF (Internet Engineering Task Force) propuso la adopción de un nuevo protocolo al que se denominó IPv6 (Internet Protocolo versión 6).

Lentamente, desde hace varios años, IPv6 está reemplazando a IPv4 en Internet mediante la aplicación de variadas técnicas de transición. Hoy coexisten en Internet ambos protocolos. El avance del despliegue de IPv6 en Internet no es uniforme a nivel mundial y existen regiones en las cuales todavía este no es importante. Este es el caso de la Argentina.

Este trabajo tiene como propósito principal aportar información que facilite e incentive el despliegue de IPv6 en redes de área local (LAN – Local Area Network) como forma de mejorar la penetración del nuevo protocolo tanto en ámbitos académicos, pequeñas organizaciones en general y en el hogar.

La poca comprensión del nuevo protocolo y su tardía adopción en nuestras redes podría traer como consecuencia elevados costos derivados de apresurados despliegues y, en el peor de los casos, pérdidas de conectividad a sitios de Internet que solo operen con IPv6.

Palabras clave: IPv6 – Transición – Coexistencia - Despliegue.

CONTEXTO

La presente investigación se enmarca en el proyecto "Estrategias para el despliegue de IPv6 en redes LAN", aprobado por el Área de Investigación y Desarrollo Científico del Departamento Académico San Salvador de la Universidad Católica de Santiago del Estero (UCSE-DASS), mediante Disposición Nº461-2018 y se encuentra financiado por esta entidad.

El equipo de investigación está constituido por docentes de las cátedras de Información y Comunicación y Redes de Computadoras de la carrera de Ingeniería en Informática, un docente asesor y cuenta con el aporte de alumnos investigadores de los últimos años de la carrera.

1. INTRODUCCIÓN

El agotamiento de las direcciones IPv4 es una realidad que afecta e involucra a los sectores público y privado, a la comunidad técnica en particular, a la sociedad civil en general, y por supuesto, a la comunidad académica, especialmente a los investigadores del área de las comunicaciones [1]. El despliegue tardío de IPv6 [2] podría traer como consecuencia un impacto negativo en la experiencia de Internet por parte de los usuarios y la potencial pérdida de negocios para los proveedores de servicio.

Una buena razón para iniciar la transición hacia IPv6 es la posibilidad de incorporar nuevos servicios en la red. Estos seguramente requerirán de innovación en el campo de la industria del software. Por otro lado, IPv6 mejora la eficiencia en la gestión de las redes facilitando la administración de la infraestructura de comunicaciones y simplificando el ruteo. Sus características además, facilitan la implementación de redes

de sensores y dispositivos IoT (Internet of Things), mejorando la escalabilidad de las redes y sus servicios [3].

Los problemas de IPv4 están relacionados con el agotamiento de sus direcciones, con problemas en la escalabilidad del ruteo y en la ruptura del esquema de comunicaciones extremo a extremo (end-to-end) originalmente diseñado [3].

El despliegue masivo de dispositivos móviles con capacidades de conectividad ha incrementado rápidamente la demanda de direcciones de red [2].

La subdivisión de bloques utilizada como forma de hacer mejor uso de las direcciones IPv4 restantes trajo como consecuencia la desagregación de prefijos, además, las técnicas de multihoming y la ingeniería de tráfico han causado un incremento en las tablas de ruteo IPv4 globales, dificultando el forwarding en internet y planteando un problema de escalabilidad de rutas [3].

El uso extendido de NAT (Network Address Translation) ha dificultado el desarrollo de aplicaciones que hagan uso de la propiedad de comunicaciones end-to-end sobre toda Internet.

IPv6 fue desarrollado como el protocolo de red de nueva generación, proponiéndose como la superación a los problemas de IPv4. Sin embargo, IPv6 no fue diseñado de manera compatible con IPv4, lo que significa que las redes IPv6 no pueden comunicarse con redes IPv4 naturalmente. Dada la incompatibilidad entre ambos protocolos, estos coexistirán durante un período más o menos prolongado y el proceso de transición será gradual.

Durante el período de transición deberá administrarse la disponibilidad de tanto IPv4 como IPv6 y resolver las cuestiones derivadas de la implementación de DNS (Domain Name System), QoS (Quality of Service), seguridad y otros aspectos abarcados por el entorno de doble pila (Dual Stack) [4].

Se necesitan un número de técnicas de transición para mantener la conectividad tanto de IPv4 como de IPv6 [5, 6].

La coexistencia de redes obligará a switches, routers, firewalls y servidores a expandir su hardware y software para soportar las características de IPv6.

El nuevo conjunto de protocolos IPv6 tales como Neighbor Discovery [7], ICMPv6 (Internet Control Message Protocol Version 6) [8] y DHCPv6 (Dynamic Host Configuration Protocol Version 6) [9] pueden hacer surgir nuevos riesgos en materia de seguridad que deberán ser evaluados.

Los hosts en los extremos requerirán cierta inteligencia para decidir cual stack de protocolos usar cuando DNS responda tanto con registros A como con registros AAAA [3].

Básicamente existen dos formas de configurar una red LAN IPv6. Uno de los métodos es la configuración manual y el otro es la configuración automática.

La configuración automática o autoconfiguración se implementa mediante el protocolo Neighbor Discovery mediante mensajes ICMPv6. Estos mensajes se denominan NS (Neighbor Solicitation), RS (Router Solicitation), NA (Neighbor Advertisement) y RA (Router Advertisement).

En la autoconfiguración el router anuncia el prefijo de red a utilizar y el host mediante EUI-64 (Extended Unique Identifier - 64) autocompleta la dirección en base a su dirección MAC (Media Access Control). También se utilizan técnicas aleatorias para autocompletar las direcciones IPv6.

Originalmente el método de autoconfiguración no proveía anuncios de servidores DNS, pero actualmente puede usarse la opción RDNSS (Recursive DNS Server) (RFC8106) en los anuncios RA [10].

También se puede hacer uso de un servidor DHCPv6. Mediante DHCPv6 no solo pueden anunciarse prefijos de red sino también otros datos como, por ejemplo, servidores DNS [11].

Para que los hosts IPv6 tengan conectividad a Internet se plantean dos posibles alternativas desde el punto de vista del servicio disponible en el ISP (Internet Service Provider). Algunos ISP proporcionan IPv6 en modo nativo, en estos casos no serán necesarias configuraciones adicionales dado que el equipo del proveedor se comunicará con el

router de borde de la LAN mediante sesiones BGP-4 (Border Gateway Protocol Version 4) o bien a través de rutas estáticas [12].
En caso de que el ISP no soporte IPv6 en forma nativa existen diversas técnicas de túnel las cuales, básicamente, encapsulan datagramas IPv6 en datagramas IPv4.
Los túneles pueden implementarse de forma manual o automática. Entre los túneles manuales pueden mencionarse los túneles 6in4 [13]. Existe también la posibilidad de automatizar túneles 6in4 mediante un servidor Tunnel-broker. Típicamente un servidor Tunnel-broker es un router con conectividad IPv6.
Algunos ejemplos de Tunnel-broker son Freenet6, HE y SixXS.
Entre los túneles automáticos pueden mencionarse diversas alternativas como 6to4, 6RD (evolución de 6to4), DS-Lite (encapsula IPv4 en IPv6) [14] y Teredo (Miredo) [15, 16].
En caso de hosts que solo soporten IPv6 y deseen comunicarse con hosts que solo soporten IPv4 se utilizan mecanismos de traducción tales como NAT64/DNS64 [17,18,19], pero solo traduciendo TCP, UDP e ICMP.
Caso contrario serán necesarios servidores ALG (Application Level Gateway).
Para solucionar el inconveniente de NAT64/DNS64, puede emplearse 464XLAT [20, 21] que consiste en realizar una traducción doble IPv4-IPv6 mediante un servidor CLAT (Customer-side translator) y una traducción IPv6-IPv4 mediante un servidor PLAT (provider-side translator).
En caso de hosts que únicamente manejen IPv4 y deseen comunicarse con redes IPv6 no se recomiendan mecanismos de traducción sino la implementación de doble pila (dual-stack) en el sistema operativo del host solo IPv4.
El tráfico IPv6 continúa creciendo de manera gradual en la actualidad y la necesidad de realizar la transición hacia el nuevo protocolo será solo cuestión de tiempo. Lo importante en cada caso sería planificar adecuadamente los tiempos necesarios y los mecanismos de transición más adecuados.
No realizar la transición podría implicar deterioro en la experiencia de Internet, pérdida de competitividad en los negocios o, en el extremo, la pérdida de conectividad total a Internet.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

IPv6 no es un protocolo de reciente aparición, pero su adopción resulta lenta, especialmente en la Argentina. Promover el conocimiento de IPv6, capacitar a profesionales del área de las comunicaciones y fomentar su despliegue en nuestras redes constituyen tareas fundamentales para no detener el crecimiento de Internet.

Para lograr esto se proyecta:

- Realizar seminarios de difusión del nuevo Protocolo IPv6 y dar a conocer las razones del cambio de protocolo de comunicaciones.
- Brindar asesoramiento técnico para proyectos de despliegue de IPv6 en redes gubernamentales o de organizaciones públicas.
- Realizar laboratorios de configuración de dispositivos conectados mediante IPv6 y prácticas de acceso a la Internet IPv6 tanto en las cátedras relacionadas a este conocimiento como público en general.

Las instituciones académicas, especialmente las Universidades, deben contribuir en el proceso de crecimiento de Internet dado que se trata de una herramienta fundamental para el desarrollo del país.

La UCSE-DASS debe brindar estos conocimientos tanto a sus alumnos de Ingeniería en Informática como a profesionales del área de las comunicaciones, promoviendo así la rápida adopción de IPv6 no solo en ámbitos académicos sino también en las redes de organizaciones de todo tipo y en el hogar.

3. OBJETIVOS

Este proyecto tiene estipulados, los siguientes objetivos principales:

- Analizar las limitaciones actuales de IPv4 y el impacto del agotamiento de sus recursos.

- Evaluar el estado actual de las técnicas para la transición y coexistencia de IPv4 e IPv6.
- Configurar pequeñas redes LAN IPv6.
- Simular IPv6 en redes LAN.
- Estudiar el estado actual del despliegue de redes IPv6 en la Argentina.
- Evaluar los resultados y obtener conclusiones.

4. ESTRUCTURA DEL EQUIPO DE INVESTIGACIÓN

Director: Ing. Germán Eduardo Jerez. Profesor Adjunto Redes de Computadoras.
Docentes Investigadores:
- Ing. Víctor José López. Profesor JTP Información y Comunicación.
- Ing. Víctor Manuel Longo. Profesor JTP Redes de Computadoras.

Asesor del proyecto: Mg. María del Pilar Gálvez
Alumnos investigadores de la carrera de Ingeniería en Informática:
- Marcelo Alejandro Gallardo.
- Raúl Alfredo Zamora.

5. BIBLIOGRAFÍA

[1] Internet Society. *The Internet is for Everyone – IPv6 Deployment: State of play and the way forward*. ISOC, 2010.

[2] Deering, S. E., Hinden, R.,Internet *Protocol, Version 6 (IPv6) Specification. RFC 8200*. IETF, 2017.

[3] Wu, P., Cui, Y., Wu, J., Liu, J., Metz, C.. *Transition from IPv4 to IPv6: A State-of-the-Art*. IEEE Communications Surveys Tutorials, 2012.

[4]Gilligan, R. E., Nordmark, E.. *Basic Transition Mechanisms for IPv6 Hosts and Routers. RFC 4213.* IETF, 2005.

[5]Arkko, J., Baker, F.. *Guidelines for Using IPv6 Transition Mechanisms during IPv6 Deployment. RFC 6180*. IETF, 2011.

[6] China Telecom. *IPv6 Deployment Best Practice by China Telecom*. 2015.

[7]Simpson, W. A., Narten, T., Nordmark, E., Soliman, H.. *Neighbor Discovery for IP version 6 (IPv6). RFC 4861*. IETF, 2007.

[8] Gupta, M., Conta, A.. *Internet Control Message Protocol (ICMPv6) for the Internet Protocol Version 6 (IPv6) Specification. RFC* 4443. IETF, 2006.

[9] Volz, B.. *Dynamic Host Configuration Protocol for IPv6 (DHCPv6) Relay Agent Subscriber-ID Option. RFC* 4580. IETF, 2006.

[10] Jeong, J. P., Park, S. D., Beloeil, L., Madanapalli, S.. *IPv6 Router Advertisement Options for DNS Configuration. RFC 8106.* IETF, 2017.

[11] Cicileo, G., Gagliano, R., O'Flaherty, C., Olvera Morales, C., Palet Martínez, J., Rocha, M., Vives Martínez, A.. *IPv6 Para Todos.* Internet Society, 2010.

[12] Acosta, A., Aggio, S., Cicileo, G., Lynch, T., Moreiras, A. M., Rocha, M., Servin, A., Berenguer, S. S.. *IPv6 para Operadores de Red*. Internet Society, 2014.

[13] Despres, R.. *IPv6 Rapid Deployment on IPv4 Infrastructures (6rd). RFC 5569.* IETF, 2010.

[14] Durand, A., Droms, R., Lee, Y., Woodyatt, J.. *Dual-Stack Lite Broadband Deployments Following IPv4 Exhaustion. RFC 6333*. IETF, 2011.

[15] Huitema, C.. *Teredo: Tunneling IPv6 over UDP through Network Address Translations (NATs). RFC 4380.* IETF, 2006.

[16] Koršič, L., Istenič, M. S.. *IPv4/IPv6 Transition Mechanisms*. IETF/RIPE, 2011.

[17] Latin American and Caribbean Internet Addresses Registry. *Implementando NAT64 / DNS64*. LACNIC, 2011.

[18] Matthews, P., Van Beijnum, I., Bagnulo, M.. *Stateful NAT64: Network Address and Protocol Translation from IPv6 Clients to IPv4 Servers. RFC 6146*. IETF, 2011.

[19] Matthews, P., Sullivan, A., Van Beijnum, I., Bagnulo, M.. *DNS64: DNS Extensions for Network Address Translation from IPv6 Clients to IPv4 Servers. RFC 6147.* IETF, 2011.

[20] Palet Martínez, J.. *464XLAT. LACNIC23*, Lima (Perú), 2013. Internet Engineering Task Force (IETF).

[21] Mawatari, M., Kawashima, M., Byrne, C..*464XLAT: Combination of Stateful and Stateless Translation. RFC 6877.*IETF, 2013.

Tecnologías de Comunicaciones para Zonas con poca Densidad Poblacional

Antonio Castro Lechtaler [1, 2, 3 y 4]; Alejandro Arroyo Arzubi [1]; Fernanda Carmona [3;]
Antonio Foti [4]; Alberto Riba [3], Rafael Mario Olivieri [1]; Germán Kurt Grin [2], Fernando Vera Batista [1].

[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; albertoriba@gmail.com; rmolivieri@gmail.com; german.grin@gmail.com; verabatista@est.iue.edu.ar.

RESUMEN

En la Argentina, como en otros países en vías de desarrollo, muchas regiones rurales con baja densidad poblacional carecen de servicios de comunicaciones por falta de interés comercial de las empresas públicas de telecomunicaciones en brindar estas prestaciones.

Esta situación dificulta a la población rural acceder a una adecuada educación, a la actividad económica, la atención de la salud en casos de urgencia, y provoca todo tipo de problemas a las comunidades que tienen estas carencias. El reciente crecimiento de la teledensidad[1] en las zonas urbanas, impulsado por la tecnología móvil, ha hecho que la brecha digital entre las zonas rurales y urbanas se haya ampliado notablemente [1].

Este Grupo de Investigación trabaja estudiando 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 o en otras de mayor tamaño 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.

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.

Los trabajos que se realizan con el grupo de investigadores tienen como objetivo principal hallar soluciones técnicas, a costos razonables[2], para los problemas de conectividad descriptos.

Palabras Clave:

CSMA/CA, WLAN, 802.11, 802.22, TVWS.

CONTEXTO

El Grupo tiene su sede en el Laboratorio de Redes – ***RedLab**, de* la Escuela Superior Técnica "Gral. Div. Manuel N. Savio" (EST)

[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.

[2] Pues si bien las soluciones técnicas existen, en la mayoría de los casos sus costos de operación hacen muy difícil su sostenimiento por parte de los usuarios.

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), mediante un trabajo en Red de Universidades nacionales.

Hasta el momento los trabajos realizados 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 púbico con jurados externos, la Universidad de la Defensa Nacional otorgó a este proyecto un Subsidio Especial [3] de 120.000,00 $ para continuar con su ejecución durante el año 2019. Y el resto de las Universidades participantes están también analizando los montos que asignarán a este proyecto.

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, 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.

Pensamos que estas nuevas tecnologías podrían ser la base para posibilitar obtener la conectividad a estas zonas apartadas, funcionando armónicamente con los canales de televisión, dándole a éstos además, valor agregado a este servicio que ya está instalado y funcionando.

1. INTRODUCCIÓN.

A causa del alto grado de saturación que presenta el espectro de frecuencias, se está trabajando en nuevas tecnologías que permitan su reutilización. Para ello, se han ido desarrollando nuevas tecnologías más acordes al estado actual del arte, como es el caso de los equipos ya existentes en el mercado que cumplen con la Recomendación 802.22, de la IEEE [4].

Las sociedades modernas están confiando y utilizando cada vez más el espectro radioeléctrico hasta el límite. Ello ha provocado que en el caso de la telefonía se esté produciendo una constante disminución de la cantidad de líneas por habitantes en todos los países del mundo.

En su reemplazó aparecen masivamente las tecnologías inalámbricas que dan lugar a distintos servicios de comunicaciones, a través de nuevos dispositivos, tales como telefonía móvil, transmisión de audio y video en tiempo real (streaming), audio y video conferencias, Wi-Fi, radiodifusión de Televisión Digital Abierta de Alta Definición, uso de la nube (entre otros). Así, el espectro de frecuencias se ha convertido en uno de los recursos más necesarios y escasos de los tiempos modernos.

Por otra parte, el crecimiento de la demanda mundial para el tráfico de datos móviles se ha incrementado en los últimos años a tasas de crecimiento anual cercanas al 70 %, y en muchos casos aún mayores.

De acuerdo con el Índice de Red Visual de Cisco (VNI) de Pronóstico del Tráfico de Datos Móviles Globales en el periodo 2011 al 2016, el tráfico de datos móviles se estimaba que iba a aumentar 18 veces en los próximos cinco años, alcanzando los 10.8 Exabytes por mes, a una tasa anual de 130 Exabytes para el año 2016.

El fuerte aumento esperado en el tráfico móvil se debía, en parte, al incremento proyectado del número de dispositivos conectados con Internet Móvil, que iba a exceder el número de personas que habitan sobre la tierra (ya se estimaba que la población mundial para el año 2016 iba a ser de 7.3 mil millones)[3]; y para el periodo 2011−2016, se pronosticaba que el tráfico de datos móviles globales crecería tres veces más que el tráfico de datos fijos globales.

[3] Fuente: Naciones Unidas.

La realidad resultó otra cosa. El tráfico móvil a nivel global y en América Latina en particular, creció en el periodo 2010 - 2018 de manera mucho mayor a los valores esperados. Alcanzó durante 2018, un valor 57 veces mayor, que la cantidad total del tráfico de datos móviles del año 2010, dato éste que resultó mucho mayor al esperado.

El crecimiento es no solo un indicativo de que la movilidad se está convirtiendo en una característica crítica de casi toda experiencia de red, sino también del valor que los consumidores y negocios le otorgan, lo que representa enormes oportunidades por delante para los proveedores de servicio que estén en el centro de la Red Internet. El video y la nube, son los grandes protagonistas" [5[.

Resumiendo: En el año 2018 [6] solamente en América Latina, el crecimiento de datos móviles creció, generando casi 1,2 Exabytes por mes, y a nivel global, aumentó casi 11 veces, alcanzando una tasa anual de 190 Exabytes; lo que es el equivalente a tener 42 billones (42 x 1012) imágenes. Esto significa en términos prácticos, unas 15 imágenes diarias por cada persona que vive en el planeta durante un año (tenga o no servicios móviles).

Para Cisco, el enorme incremento durante el año pasado estuvo impulsado por una serie de factores, tales como:

- El continuo crecimiento del número de conexiones de Internet móvil, especialmente desde dispositivos personales.
- Las conexiones de máquina a máquina (M2M), los cuales excedieron los 10 mil millones.
- El aumento en las velocidades móviles promedio, que han pasado de 1.4 Mbps en 2013 a 2.5 Mbps en 2018,
- Y a un mayor consumo de video móvil en dispositivos inteligentes, que llegó a representar el 69% del tráfico de datos móvil global.

A nivel de la región, se proyecta que el aumento del tráfico de datos móviles alcance durante el corriente año 2019, 1.500 Exabytes[4], distribuidos de la siguiente manera:

- Argentina: 110 Petabytes.
- Brasil: 490 Petabytes.
- Chile: 70 Petabytes.
- México: 290 Petabytes.
- Resto de América Latina: 390 Petabytes.

Por otra parte, el uso del espectro de hasta 10 GHz y a futuro a frecuencias aún más altas, ha dado lugar a una revisión de las políticas de regulación del mismo, a nivel de la Unión Internacional de Telecomunicaciones - ITU.

A causa del congestionamiento en esas frecuencias, se ha intensificado el estudio y las aplicaciones que permitan utilizar adecuadamente, los llamados Espacios Blancos - White Space, que se consideran muy importantes en la ejecución de este trabajo.

La Conferencia Europea de Administraciones de Correos y Telecomunicaciones - CEPT, ha definido a un Espacio Blanco como: "*Una porción del espectro que está disponible para ser utilizado por una aplicación de radiocomunicaciones en un momento dado, para una zona geográfica determinada, en forma simultánea con otra, pero con la característica de que la misma no produzca interferencias con respecto a los servicios que poseen una prioridad más alta a nivel nacional en esas mismas frecuencias*"[7].

En la actualidad, son intensos los esfuerzos de investigación realizados a nivel de organizaciones no gubernamentales, países y empresas para darle utilidad a esta parte del espectro que es desaprovechado.

La importancia del tema, ha dado lugar a la creación de la "WhiteSpace Alliance[5]"cuya misión es "Promover el desarrollo, despliegue y uso de estándares basados en productos y servicios como un medio para proporcionar capacidades de banda ancha a través de los espacios blancos existentes en el espectro de frecuencias".

Dada la inversión y el fuerte despliegue que la Republica Argentina ha efectuado al crear e instalar el sistema de Televisión Digital

[4] 1 Exabyte = 10^{18} Bytes

[5] https://www.whitespacealliance.org/

Abierta - TDA, el mismo (infraestructura ya existente) ofrece una inmejorable oportunidad para solucionar el problema de las comunicaciones rurales o de poblaciones aisladas o con baja densidad poblacional utilizando los espacios blancos que ese sistema permite reutilizar

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. También, sus resultados pueden ser útiles para las comunicaciones militares en ambientes en los que los sistemas fijos son inexistentes o precarios. Este proyecto puede ser considerado dual: civil y militar.

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 que permitiría el funcionamiento de los equipos que brindan servicios de comunicaciones rurales.

2.5. Estudio práctico y teórico, de las interferencias entre canales debido al uso intensivo de las comunicaciones inalámbricas para todo tipo de servicios de comunicaciones [8].

2.6. Determinación de la posible utilización de las frecuencias asignadas para la transmisión del Sistema Nacional de Televisión Abierta, para combinarlo con un sistema basado en estas normas y así emplearlo 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. Hasta el momento 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 [9, 10] 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 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.

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 que ellos podrían 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 [11] 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 (solamente 31 trabajos sobre el total presentado con referato) [12], 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 [13].

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 2017 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 2018 y en este año 2019, se han sumado al proyecto nuevos investigadores noveles 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 que algunos de ellos, realicen su Trabajo Final de Carrera en algún tema de los que aborda la presente línea de investigación.

Los integrantes de este Equipo de Investigación multidisciplinario son docentes y alumnos de las siguientes asignaturas en las distintas Universidades participantes. En particular de las asignaturas: 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 Nº 309/2018 de fecha 14 de diciembre de 2018. Expediente Nº 163/2018.

[4] IEEE 802.22 - Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Policies and Procedures for Operation in the TV Bands.

[5]https://searchdatacenter.techtarget.com/es/cronica/Trafico-de-datos-moviles-en-AL-aumentara-en-casi-13-veces-para-2018.

[6]https://colombiadigital.net/actualidad/noticias/item/9520-el-trafico-mundial-de-datos-aumentara-7-veces-entre-2016-y-2021.html.

[7] 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.

[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 Regional Area Network - Medium Access Control (MAC) and Physical Layer (PHY). Specifications: Policies and Procedures for Operation in the TV Band.

[10] LAN: Local Area Network; MAN: Metropolitan Area Network.

[11] 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.

[12] 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.

[13] A. Castro Lechtaler, A. Foti, R. Fusario, C. García Garino and J. García Guibout. 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. pp.1.117 a 1.126. Jujuy. October 2009.

Uso colaborativo del procesador en sistemas embebidos para múltiples interfaces

Waldo Valiente, Esteban Carnuccio, Mariano Volker, Graciela De Luca, Raúl Villca, Germán Lorenz
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Dirección: Florencio Varela 1703 – CP 1754 –
{wvaliente, ecarnuccio, mvolker, gdeluca }@unlam.edu.ar,
{raul.villcasd, germangelv}@gmail.com

RESUMEN

Las soluciones IoT requieren de sistemas embebidos cada vez más complejos, como es el caso de nuestra investigación para el monitoreo de personas, asistencia ante caídas, signos vitales anómalos y seguridad. Donde muchas interfaces entre los sensores y los canales de comunicación se deben atender junto con la lógica del programa principal. Estos a su vez se construyen sobre pequeñas arquitecturas con procesadores de rendimiento moderado impulsados por el bajo consumo energético que se busca. En el presente trabajo se analizan técnicas de programación alternativas y funcionalidades provistas por los Sistemas Operativos de Tiempo Real, que brindan la posibilidad de alternancia de tareas sobre un único procesador. También al mismo tiempo se evalúan las necesidades de recursos, que en estos tipos de procesadores son reducidas. Se proponen soluciones para compartir el procesador evitando en lo posible el uso de funciones provistas por los sistemas operativos de tiempo real, ya que estos consumen una cantidad de recursos considerables para estas arquitecturas. Se propone solucionar esta dificultad mediante la implementación de una máquina de estados a través del uso de interrupciones, como única solución o una combinación de ambas.

Palabras clave: *RTOS, interrupciones, monitoreo de personas, máquina de estado, funciones bloqueantes.*

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 CYTMA2 de la Universidad Nacional de La Matanza, el cual está 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 e Internet de las cosas (IoT), entre otros.

1. INTRODUCCIÓN

Los sistemas embebidos que forman parte de un sistema complejo de IoT [1], se construyen sobre arquitecturas simples y colaborativas como el procesador ARM de la familia Cortex-M3 [2] [3]. Estos sistemas tienen la complejidad de atender y gestionar los recursos para múltiples interfaces sobre un único hilo de ejecución. La dificultad radica en cómo repartir el recurso escaso, en este caso el procesador, brindando atención a cada interfaz, siempre manteniendo los tiempos de respuesta requeridos en todos los dispositivos que se utilizan. En nuestra investigación [4] se evidencia que cuando se utilizan interfaces complejas se requiere realizar una serie de pasos o procedimientos compuestos para lograr dar respuesta a una solicitud. Como puede ser el caso de conexión a Internet por GSM, conexión local por Bluetooth y obtener datos desde el GPS, junto a otros sensores que requieren un muestreo periódico. Tal es el caso del acelerómetro, donde se evalúan las mediciones para verificar si la persona que lleva el dispositivo se encuentra en una caída o no.

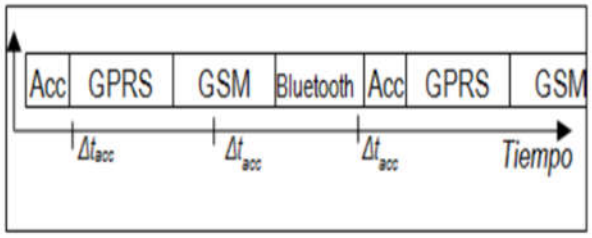


Figura 1 – Múltiples funciones sobre un único hilo de ejecución.

Se puede observar en la Figura 1 que sobre un único procesador ejecutan varias funciones. La función que requiere ejecutar regularmente y la del acelerómetro (Acc), no siempre logra cumplir con su periodicidad (Δt_{acc}) requerida para el correcto funcionamiento del algoritmo de caídas.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación se plantean distintos métodos para lograr repartir el uso del único procesador entre varias funciones que se necesitan ejecutar del sistema embebido y reconociendo aquellas que son críticas.

Las implementaciones en sistemas embebidos podrían clasificarse en non-os, bare-bones con RTOS [1] y Full RTOS. En el caso de una implementación sin S.O se programa un binario con todo el programa en sí, que realiza la inicialización-ejecución principal y finalización. En el caso de bare-bones con RTOS es un mix entre ambos métodos. Por último en el caso del uso de Full RTOS, la funcionalidad necesaria se implementa como una aplicación convencional en un sistema operativo de tiempo real, Figura 2.

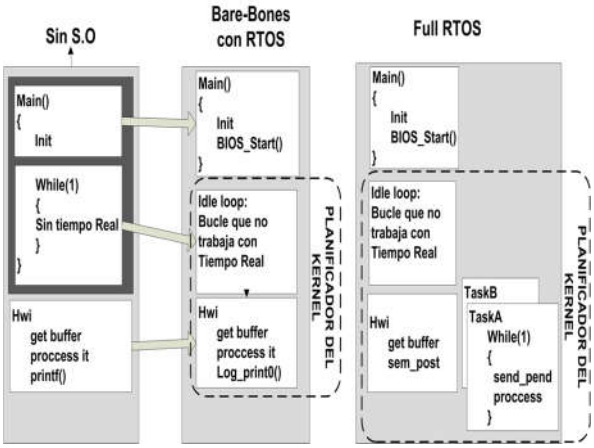


Figura 2 –Clasificación de implementación en sistemas embebidos.

Cambios de contexto por hardware

La arquitectura ARM Cortex-M3 [5] facilita dos mecanismos para interrumpir el programa en ejecución. El primero *SysTick* es un reloj de 24 bits que puede interrumpir al procesador cada vez que su contador llegue a cero. Este mecanismo es utilizado para realizar cambios de

[1] RTOS Real Time Operating System

contexto al estilo *round-robin*, Figura 3.

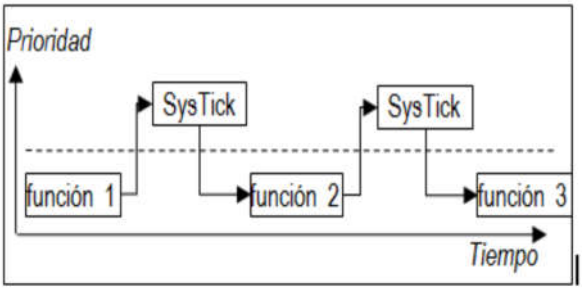


Figura 3 - Interrupción SysTick.

En el segundo mecanismo *PendSV* el programa permite ceder el control del CPU al manejador, cuando este se encuentre inactivo. Se puede utilizar con cambios de contexto al estilo de FIFO Figura 4.

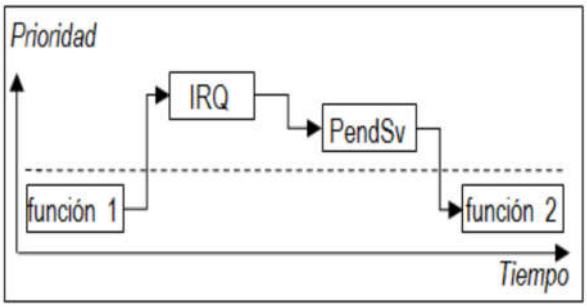


Figura 4 - Interrupción PendSV.

La ventaja de este método es que el programador tiene el control y puede definir el comportamiento que necesite para su implementación. La desventaja se tiene, cuando debe desarrollarse la lógica del planificador en forma más robusta, esta tarea suele ser engorrosa y difícil de probar la integridad del mismo.

Cambios de contexto por software

Los cambios de contexto se realizan agregando bibliotecas de terceros en el programa. Estas poseen tal nivel de integración brindando servicios directamente a nivel de Sistema Operativo de Tiempo Real (RTOS) [6] en reducido tamaño, que facilitan su programación e implementación. En el mercado existen abundante cantidad de desarrollos para la arquitectura en cuestión, a continuación se enumeran los más destacadas:

La versión de *mbedOS* [7], entre sus principales características se destaca la seguridad y la conexión de dispositivos, mientras que *FreeRTOS* [8], facilita la programación, implementación e integración con servicios en la nube de Amazon™. El caso de *NuttX* [9], pensado para dispositivos desde 8 a 32bits basado en el estándar de APIs Unix. El pequeño kernel de *µ-velOSity* [10], es diseñado para ser eficiente y ocupar poca memoria. *RTEMS* [11], es soportado en gran cantidad de arquitecturas. *ChibiOS* [12], es un pequeño kernel que permite la integración con otros componentes *OpenSource* para brindar más servicios.

Los RTOS facilitan la implementación ya que los cambios de contexto, junto con otras funciones de comunicación de tareas, ya se encuentran integradas en el kernel, el desarrollador solo se limita a desarrollar la lógica de la aplicación para las tareas que ejecutan ver Figura 5.

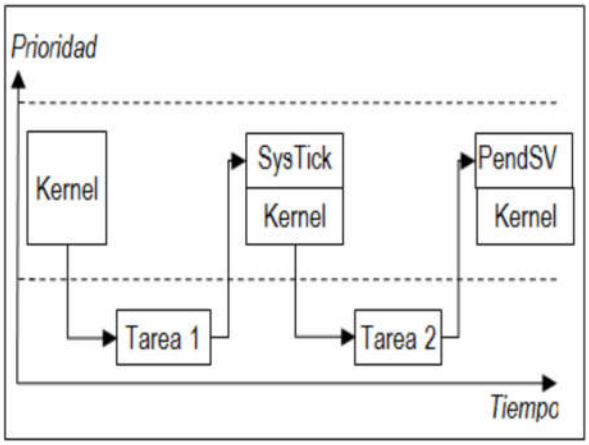


Figura 5 – Cambio de contexto utilizando RTOS

La ventaja de este mecanismo es que soluciona

los problemas planteados con bibliotecas probadas por la comunidad y acelera el tiempo de desarrollo. La desventaja que es dificultosa la coordinación entre tareas fuertemente acopladas y la utilización de software de terceros amplía el espacio que ocupara el programa en el reducido espacio de memoria que posee la arquitectura embebida.

Modelo máquina de estado finito

La alternativa que plantea nuestro grupo de investigación, permite distribuir el uso del único procesador entre varias funciones. La misma radica en descomponer la lógica de cada función en sub-partes no bloqueantes. Así el programa principal, compagina su ejecución entre las distintas sub-partes de cada función.

El método encontrado consiste en organizar esta división utilizando el modelo de máquina de estado finito. Así las funciones de las interfaces tienen un comportamiento distinto, según el estado que se encuentre. En cada interacción del programa se evalúa el estado en el que se encuentra la aplicación, ejecutando la pequeña sub-parte que le corresponde. A continuación, se evalúa la siguiente porción de código de la función siguiente.

En la Figura 6, se observa que la función del acelerómetro (Acc) logra cumplir con la periodicidad requerida. Mientras se altera la ejecución de las sub-partes del GPRS (G_{e1},G_{e2},G_{e3},G_{e4}), para el GSM (S_{e1},S_{e2},S_{e3}) y Bluetooth (B_{e1},B_{e2}).

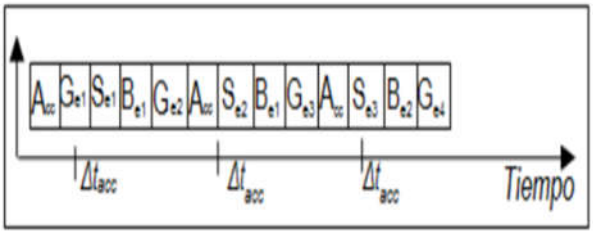


Figura 6 – Alternancia por estado de funciones.

La ventaja de esta solución es que no ocupa mucho lugar en memoria y no requiere de grandes cambios de implementación, comparándolo con la migración del programa al utilizar RTOS. Como desventaja, sin un buen diseño de la transición de estados, puede haber partes de funciones que nunca se ejecuten. Otra desventaja es que no se diferencian prioridades entre las funciones de interfaces.

3. RESULTADOS OBTENIDOS/ESPERADOS

Luego del análisis de las alternativas se confirma, que se puede obviar la utilización de un RTOS siempre que el programa no utilice funciones bloqueantes. Los RTOS son una solución compleja que brinda beneficios en casos que el código sea bloqueante, sin poder desarrollarlo de otra forma. La propuesta de máquina de estado es equitativa en el uso del procesador, pero no satisface estrictamente el cumplimiento de la periodicidad necesaria en la función crítica. En la investigación en curso se realiza la prueba de una solución hibrida. La lógica de las interfaces de GSM, GPRS y Bluetooth, menos prioritarias, se desarrollarán siguiendo el modelo de máquina de estado finito, que permite el uso del procesador colaborativamente.

Para la interfaz de acelerómetro, que es pieza clave en la detección de caídas. Se realizara un pequeño módulo que permita cambios de contexto utilizando interrupciones *Systick*. Luego de tener las mediciones periódicamente, se evaluará el algoritmo de caídas junto con las demás interfaces antes mencionadas.

4. FORMACIÓN DE RECURSOS HUMANOS

La presente línea de investigación dentro del departamento de Ingeniería e Investigaciones

Tecnológicas forma parte del trabajo que dos investigadores se encuentran realizando su tesis de maestrías. Completan el grupo de investigación cuatro alumnos de Ingeniería en Informática que se encuentran finalizando su formación de grado y realizan su iniciación a la investigación.

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] STMicroelectronics, «Referece Manual - STM32F101xx, STM32F102xx, STM32F103xx, STM32F105xx and STM32F107xx advanced Arm®-based 32-bit MCUs,» STMicroelectronics, 2018.

[3] STMicroelectronics, «Datasheet - production data Ultra-low-power 32-bit MCU ARM®-based Cortex®-M3,» de *STM32L151x6/8/B STM32L152x6/8/B*, STMicroelectronics, 2016.

[4] E. C. M. V. G. D. L. G. G. D. G. R. V. M. V. Waldo Valiente, «Dispositivo de asistencia de personas mediante monitoreo IoT,» SEDICI, San Luis, 2018.

[5] T. Martin, «The Designer's Guide to the Cortex-M Processor Family,» India, elseiver, 2016, pp. 71-130.

[6] L. P. L. P. M. T. Fei Guan, «Open source FreeRTOS as a case study in real-time operating system evolution,» de *Journal of Systems and Software*, elsiever, 2016, pp. 19-35.

[7] A. Mbed, «Mbed OS,» 2018. [En línea]. Available: https://www.mbed.com/en/platform/mbed-os/.

[8] Amazon Web Services, «Amazon FreeRTOS Sistema operativo compatible con IoT para microcontroladores,» 2019. [En línea]. Available: https://aws.amazon.com/es/freertos/.

[9] G. Nutt, «NuttX Real-Time Operating System,» 2019. [En línea]. Available: https://nuttx.org/.

[10] Green Hills Software, «µ-velOSity real-time operating system (RTOS),» 2019. [En línea]. Available: https://www.ghs.com/products/micro_velosity.html.

[11] On-Line Applications Research (OAR) Corporation, «RTEMS - An Open Real-Time Operating System,» 2019. [En línea]. Available: http://rtems.com/.

[12] G. D. Sirio, «ChibiOS free embebbed RTOs - ChibiOS Homepage,» 2017. [En línea]. Available: http://www.chibios.org/dokuwiki/doku.php.

[13] STMicroelectronics, «Datasheet - production data- Medium-density performance line ARM®,» de *STM32F103x8*, STMicroelectronics, 2015.

Virtualización de funciones de red

Peliza Carlos, Dufour Fernando, Serra Ariel, Micieli Gustavo, Guerrero Facundo
Departamento de Informática e Investigaciones Tecnológicas
Universidad Nacional de La Matanza
Florencio Varela 1908, San Justo
pelizac@gmail.com; fdufourf@hotmail.com ; aserra@unlam.edu.ar ; gmicieli@gmail.com

RESUMEN

Dentro de lo que se denomina drivers en la industria de Telecomunicaciones, o sea factores que se analizan para adoptar una tecnología novedosa, en el año 2012, durante el Congreso "SDN and OpenFlow World Congress" del mes de Octubre en Darmstadt, Alemania fue presentado el whitepaper "Network Functions Virtualization" con el objetivo de describir los beneficios, habilitadores y desafíos para la virtualización de funciones de red y como justificación para fomentar una colaboración internacional que acelerara el desarrollo y despliegue de soluciones interoperables para servidores de alto volumen, basadas en estándares.

En 2018, la consultora IDC, en su informe - Market Perspective - Doc # US43696118, nos dice "Cuando llegue el 5G comercial, las partes interesadas de la industria que virtualizaron y controlan la mayoría de la infraestructura se posicionarán para superar el desempeño. Esto será evidente cuando se implemente el núcleo 5G basado en estándares, y los proveedores de servicios de comunicación comiencen a aprovechar la división de la red. Aquellos que pueden para administrar y entregar de manera centralizada segmentos de red en toda la pila de red, incluidos los niveles de transporte y RAN, es probable que obtenga una ventaja competitiva tanto en costo como en agilidad de servicio sobre sus pares ", dice Patrick Filkins, analista de investigación senior, IoT y Mobile Network Infraestructura.

Transcurridos 6 años de la presentación cual es el estado de la virtualización, ¿se ha estandarizado?, ¿se ha extendido su uso?, ¿es una arquitectura masiva?

El presente trabajo tiene como objetivo analizar la arquitectura NFV y distinguir en ella los desarrollos que son posibles de seguir en nuestro país y la región, verificar cuál es la expansión de la arquitectura NFV.

El estilo de este trabajo de investigación es comparativo y se basa en el análisis de fuentes bibliográficas y documentación existente, junto a la realización de pruebas de concepto y funcionamiento de posibles servicios virtualizados o a virtualizar durante la investigación.

Palabras Clave: Virtualización de funciones de red, Telecomunicaciones, redes definidas por software

CONTEXTO

La secretaria de investigaciones tecnológicas, en conjunto con del Departamento de Ingeniería e investigaciones tecnológicas, son quienes dictan las políticas de investigación y las entienden como la base de producción y distribución de conocimiento. Particularmente este grupo de investigación ha venido desarrollando una serie de investigaciones orientadas a las redes de comunicaciones o a las telecomunicaciones en general, como antecedentes se pueden citar: C164 (Carrier- Ethernet), C189 (LTE) y trabajos en diferentes congresos como "Carrier Agreggation", "Narrow Band IoT", "VoLTE".

Temáticas que luego son introducidas en cátedras de comunicaciones de la Universidad Nacional de La Matanza, en carreras de Informática como de Electrónica. Adicionalmente, se brindan charlas de divulgación para la comunidad educativa de acuerdo a las necesidades del centro de estudiantes de dicha institución

1. INTRODUCCION

La virtualización de las funciones de red (NFV) es un enfoque de red propuesto por ETSI (European Telecommunications Standards Institute) que permite la sustitución de dispositivos de hardware dedicado, tales como routers, firewalls y equilibradores de carga entre otros equipamientos, por dispositivos basados en software que se ejecutan como máquinas virtuales en servidores estándares de la industria.

El viaje a una red NFV completamente operacional requiere la coordinación de tres desarrollos interconectados, pero totalmente separados: Virtualización, Orquestación y Automatización, ninguno de estos caminos puede considerarse de manera aislada (Ashton Metzler, 2015, p.47)

NFV desacopla las funciones de la red de dispositivos de hardware dedicados y las traslada a uno o varios servidores virtuales, que pueden cumplir múltiples funciones en un único servidor físico. Este enfoque reduce los costos y minimiza el mantenimiento, debido a que los dispositivos virtuales reemplazan dispositivos de red basados en hardware dedicado. La NFV no debe confundirse con una red virtualizada, porque la NFV, es un marco o arquitectura que busca descargar sólo las funciones de red, y no toda la red. También es importante señalar que la NFV es diferente de una función virtual de red (VNF), un término comúnmente utilizado para describir una función de red que se ejecuta en el software de una máquina virtual (Terminology for Main Concepts in NFV, gs_NFV003v010201p).

La amplia gama de opciones de desarrollo que presenta NFV puede expresarse mediante los siguientes cuadros de elaboración propia:

Ilustración 1 – Implementación monolítica.

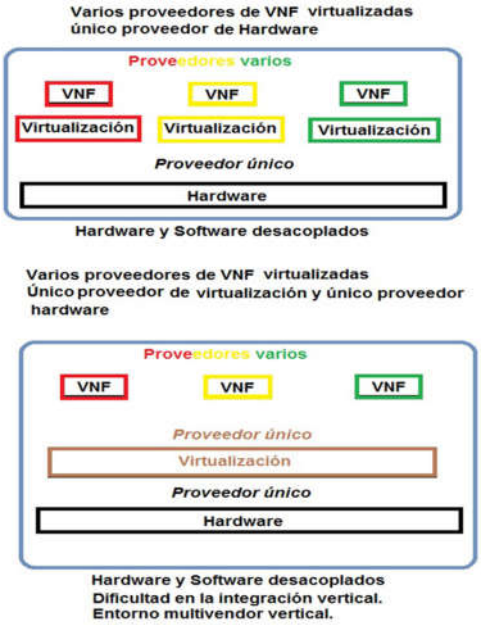


Ilustración 2 - Islas de VNF

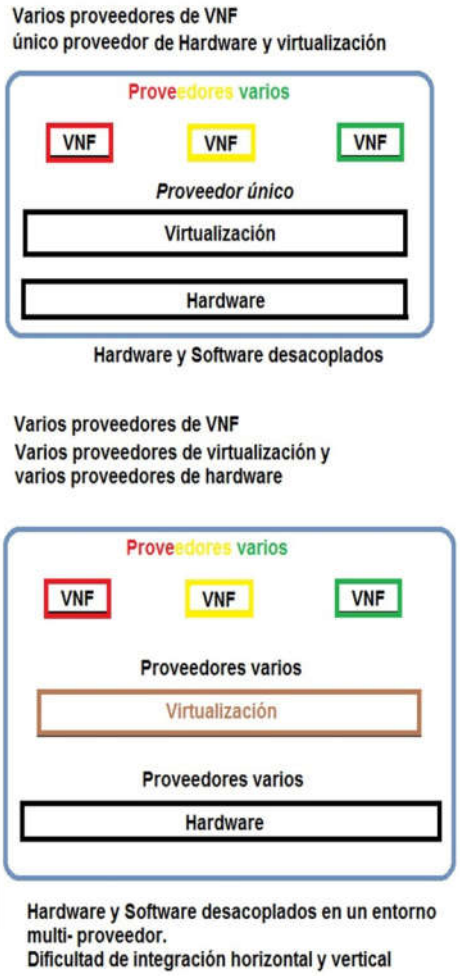


Ilustración 3 Implementación con estructura similar a SDN

En la propuesta del whitepaper del 2012 estaban definidos los campos de acción y casos de uso (NFV_White_Paper, p 6) entre los cuales se enumeraban:

- Elementos de conmutación: BNG, CG-NAT, enrutadores
- Nodos de red móvil: HLR / HSS, MME, SGSN, GGSN / PDN-GW, RNC, nodo B, eNodo B.
- Funciones contenidas en enrutadores y decodificadores domésticos.
- Análisis de tráfico: DPI, medición de QoE.
- Señalización NGN: SBCs, IMS.
- Funciones convergentes y de toda la red: servidores AAA, control de políticas.
- Optimización a nivel de aplicación: CDN, servidores de caché, equilibradores de carga.
- Funciones de seguridad: cortafuegos, escáneres de virus, sistemas de detección de intrusos, protección contra correo no deseado.

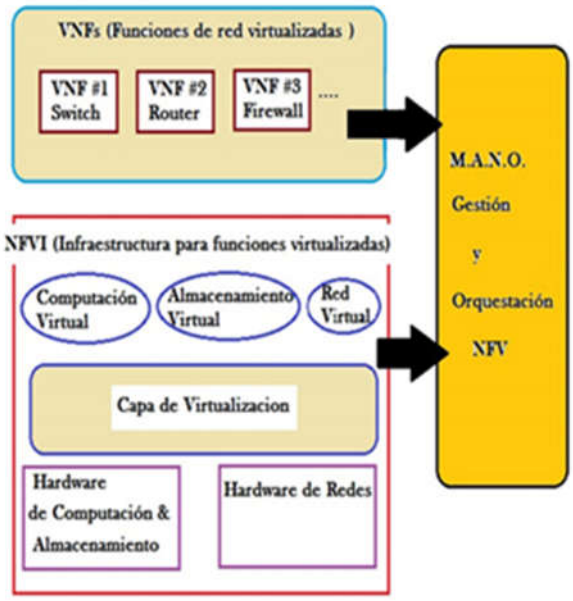


Ilustración 4 - Arquitectura NFV

Se ha comprobado la existencia de funciones simples de red virtualizadas, a saber, service border controllers, routers y switches, sin embargo ¿cuál es la actualidad de los nodos y equipos de core en redes móviles.?

2. LINEAS DE INVESTIGACION y DESARROLLO

La problemática a investigar incluye, entre otras temáticas, las percepciones de los principales especialistas en estas tecnologías y lo que ocurrirá con la misma durante los próximos 3 años.

Surgirán de los estudios distintas soluciones posibles de implementación, donde cada una de ellas llevará asociadas ventajas y limitaciones. Estas determinarán los servicios que cada caso pueda prestar.

El campo específico de la investigación consiste en trabajar sobre los aspectos de interoperabilidad y capacidades de los elementos de red de distintos proveedores, determinando su compatibilidad de funcionamiento frente a los estándares, concentrándonos en el comportamiento en este nuevo escenario tecnológico.

El modelo teórico conceptual para la problemática a analizar se estructura a partir del conocimiento de incompatibilidades históricas que presentan las tecnologías en sus estados iniciales de desarrollo con las necesidades de buen funcionamiento de los servicios.

Esto requiere un análisis de dichas incompatibilidades de acuerdo con el estado del arte, con el objeto de que los proveedores de tecnologías trabajen sobre las mismas y que los Carriers de comunicaciones tengan claro cuáles son las limitaciones a la hora de implementar sus servicios.

La validez de los resultados se obtendrá a través de la verificación de la compatibilidad de las tecnologías para los servicios convergentes y la búsqueda de pruebas exitosas que convaliden nuestras conclusiones, para lo cual una vez obtenido el resultado, se lo contrastará con pruebas de laboratorio o

pruebas piloto en operadoras de telefonía móvil que se buscarán en el mercado.

3. RESULTADOS OBTENIDOS/ESPERADOS

Obtener conocimiento e información sobre el estado de arte real de la tecnología de virtualización y sus posibilidades de implementación en Argentina.

Lograr reconocer las implementaciones más convenientes (mejores prácticas) que podrán ser utilizadas por los operadores para reducir tiempos de implantación.

Determinar una prospectiva de la evolución de este tipo de redes en los próximos 3 años.

Obtener información sobre cuáles serán los servicios que oficiarán de apalancadores del negocio de las telecomunicaciones de la región.

4. FORMACION DE RECURSOS HUMANOS

El equipo de investigadores pertenece al cuerpo docente del Departamento de Ingeniería de la Universidad de La Matanza, con miembros de las carreras de Electrónica e Informática, todos con gran experiencia en el área de telecomunicaciones.

Para el segundo año de investigación, se tiene planificado incorporar alumnos, como parte de un plan para mejorar sus habilidades blandas o soft skills, que incluyan la redacción científica de textos y/o a los que la investigación les permita formalizar su trabajo de fin de carrera.

5. BIBLIOGRAFIA

Hakiri, A., Gayraud, T., Schmidt, D., & Berthou, P. (2014). Software Defined Networking: Challenges and research opportunities for Future Internet. *Computer Networks Volume 75, Part A*, 453-471.

AT&T: Margaret Chiosi.BT: Don Clarke, Peter Willis, Andy Reid. China Mobile: Dr. Chunfeng Cui, Dr. Hui Deng.et all. (s.f.). (29 de Abril de 2018). *ETSI.* Obtenido de European Telecommunications Standards Institute: http://portal.etsi.org/NFV/NFV_White_Paper.pdf

Chris Wolf, Erick M. Halter. (2005). *Virtualization: From the Desktop to the Enterprise (Books for Professionals by Professionals).* Apress.

Metzler, A., & Metzler, J. (2015). *The 2015 guide to SDN and NFV*. Obtenido de Webtorials.

Rajendra Chayapathi, Syed F. Hassan, Paresh Shah. (2017). *Network Functions Virtualization (NFV) with a Touch of SDN.* Pearson Education.

Bases de Datos

y

Minería de Datos

Análisis comparativo de tareas de pre procesamiento de textos sobre contenido extraído de redes sociales

Leonardo Esnaola[1 4], Juan Pablo Tessore[2 4], Hugo Ramón[3 4], Claudia Russo[3 4]

Instituto de Investigación y Transferencia en Tecnología (ITT)
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

leonardo.esnaola@itt.unnoba.edu.ar / juanpablo.tessore@itt.unnoba.edu.ar / hugo.ramon@itt.unnoba.edu.ar / claudia.russo@itt.unnoba.edu.ar

Resumen

El texto que surge de la interacción entre usuarios en redes sociales suele ser más disperso que el contenido tradicional. Es decir, contiene errores ortográficos, uso informal del lenguaje, emoticones, urls y otras construcciones que no suelen estar presentes en el lenguaje formal. Dicha dispersión puede afectar el desempeño de los clasificadores de texto basados en aprendizaje automático.

El presente trabajo propone medir el desempeño de diferentes tareas de pre-procesamiento, aplicadas primero de manera aislada y luego combinadas, sobre contenido extraído de redes sociales. Se busca determinar cuán aptas resultan ser estas tareas para corregir errores en textos de este tipo. Para ello, en primer lugar, se determinará en qué magnitud se reduce el porcentaje de palabras "incorrectas" y, en segundo lugar, cómo impactan en la precisión final alcanzada por clasificadores basados en aprendizaje automático.

Este trabajo, se enmarca en una línea de investigación más amplia que propone la construcción de un clasificador automático de opiniones utilizando algoritmos de aprendizaje automático, el cual fuera presentado previamente en otra edición de este Workshop [1], y que permitirá realizar análisis automáticos de bajo costo para determinar las emociones manifestadas por consumidores o usuarios acerca de productos o servicios, a partir del análisis de sus opiniones escritas. Este clasificador será entrenado a partir de los comentarios en lenguaje informal presente en redes sociales.

***Palabras clave*:** Minería de textos, Pre-procesamiento, Inteligencia artificial, Redes sociales.

Contexto

Esta línea de investigación forma parte del proyecto "Tecnología y aplicaciones de Sistemas de Software: Innovación en procesos, productos y servicios" presentado en el marco de la convocatoria a Subsidios de

---

[1] Docente Investigador ITT / Doctorando UNLP

[2] Docente Investigador ITT / Becario Doctoral CIC

[3] Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC

[4] ITT - Centro Asociado CIC

Investigación Bianuales (SIB2019) de la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA. 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 y Transferencia en Tecnología (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.

Introducción

Con el auge de las redes sociales, la gente está cada vez más involucrada en muchos aspectos en los en los cuales antes solo eran consumidores pasivos [2]. Dichas redes permiten a las personas expresarse de manera libre y rápida acerca de una gran variedad de temas, y están siendo aprovechadas por los sitios comerciales para influenciar a los usuarios con campañas de marketing dirigidas [3]. Según [4], las reacciones de Facebook proveen una oportunidad para medir el compromiso emocional de los consumidores. Sin embargo, procesar cientos de textos para entender que emoción reflejan no es una tarea sencilla y requiere mucho trabajo manual.

Según [5], la minería de textos se encarga del descubrimiento automático o semiautomático de información nueva, previamente desconocida y de alta calidad a partir de un gran número de textos no estructurados. La información que se quiere inferir de los textos se suele especificar manualmente de antemano. En [6] los autores definen que la minería de textos combina técnicas de tres campos específicos:

- Recuperación de información: implica la recopilación de información relevante para realizar una tarea determinada.
- Procesamiento de lenguaje natural: combina una variedad de técnicas para analizar y representar textos que ocurren naturalmente en uno o más niveles de análisis lingüístico con el propósito de lograr un procesamiento de lenguaje similar al humano para una variedad de tareas o aplicaciones [7].
- Minería de datos: descubre patrones en la información estructurada que se ha construido a partir de los textos.

El desarrollo de las redes sociales ha aumentado la disponibilidad de contenido en forma de texto, creando la materia prima necesaria para aplicar allí técnicas de minería de texto y extraer información significativa.

El presente trabajo busca medir la efectividad de diferentes técnicas de pre procesamiento sobre textos "ruidosos" provenientes de las redes sociales. A diferencia de otros estudios [8], el proceso se lleva a cabo sobre un *dataset* de comentarios en español compilados desde Facebook. Teniendo en cuenta los estudios relevados, no existe una investigación que informe la efectividad de estas técnicas sobre este tipo de conjunto de datos.

Esta investigación es parte de una investigación más amplia que se centra en la

construcción de un detector automático de emociones a partir de texto donde, a diferencia de otros estudios [9], las etiquetas que denotan la emoción reflejada en cada comentario se obtienen automáticamente de las reacciones de Facebook, en lugar de clasificarlos manualmente. Esta aplicación podría ofrecer una amplia gama de aplicaciones potenciales, como detectar la emoción que surge de la opinión de grandes grupos de personas sobre ciertos productos, servicios [10, 11, 12] o incluso políticas públicas. También podría utilizarse para identificar demandas o quejas no cumplidas de ciudadanos; En seguridad, para la detección automática de factores de riesgo en redes sociales como amenazas o ciberacoso [13].

Para desarrollar esta aplicación, basándose en un enfoque de aprendizaje automático, es necesario numerizar el texto de entrada. Esto se logra utilizando algunas métricas de texto. Una de las métricas más ampliamente adoptadas es "*term frecuency - inverse document frecuency*" (tf-idf). Es bien sabido que los datos provenientes de redes sociales suelen ser muy dispersos [14], y esta dispersión puede afectar el rendimiento de las aplicaciones que se basan en estadísticas de frecuencia de palabras [15], como tf-idf.

Debido a esto, es necesario aplicar algunas tareas de pre procesamiento a los datos de entrada para mejorar el rendimiento del clasificador de emociones a implementar.

Reparar el texto de los comentarios, realizar, por ejemplo, una corrección ortográfica, eliminar signos de puntuación incorrectos, enlaces y otras intervenciones puede eliminar características innecesarias de la entrada, haciendo que el análisis posterior sea más rápido y más preciso.

Diversos trabajos que realizan pre procesamiento de textos provenientes de las redes sociales [14, 15, 16] hacen, en primer lugar, una clasificación de las palabras del *dataset* analizado en dos categorías. Estas son palabras "*in vocabulary*" (IV) y palabras "*out of vocabulary*" (OOV), siendo estas últimas las palabras incorrectas según el lenguaje formal. Para clasificar la entrada de dicha manera, se emplean un diccionario en el idioma correspondiente y una herramienta de corrección ortográfica. Entre estas últimas, las más utilizadas [17], son *Hunspell* [18] y *GNU Aspell* [19].

Una vez identificadas las OOV, se aplican una serie de filtros al texto para reducir el porcentaje de OOV sobre el número total de palabras del conjunto de datos.

Líneas de Investigación, Desarrollo e Innovación

La presente investigación se encuadra dentro del eje "Gestión de la Innovación" del mencionado proyecto SIB. Dicho eje incluye la investigación de procesos metodológicos para abordar Innovación y su implicación medible en procesos productivos.

En este sentido, es importante transformar al usuario potencial en un partícipe activo del proceso de producción de un determinado producto, o en la concepción y desarrollo de un determinado servicio. Internet permite estar en contacto con un número inmenso de potenciales usuarios, pero poder procesar todas las interacciones buscando reconocer si estos usuarios están satisfechos o no (y en qué medida) sobre alguna característica de un producto o servicio, requiere analizar estas múltiples interacciones lo cual podría consumir mucho tiempo de trabajo.

Un clasificador automático de emociones a partir del análisis de texto podría ser útil para este propósito. Así, se presentan en la Figura 1 las diferentes etapas identificadas para su desarrollo y puede apreciarse en cuál de ellas se enmarcan las tareas de pre-procesamiento presentadas en este trabajo:

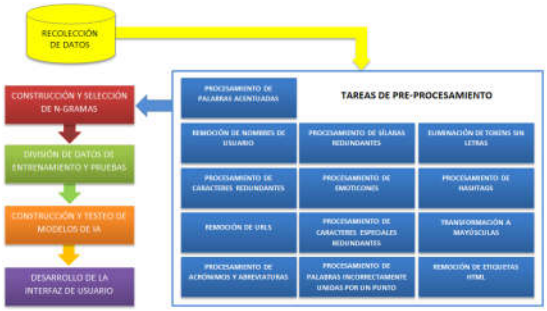


Figura 1: Etapas del proceso de construcción de la herramienta, incluyendo las tareas de pre procesamiento.

Resultados y Objetivos

Se espera que la presente línea de I/D posibilite una mejora en la exactitud de un clasificador automático de textos, a partir de la reparación de los comentarios utilizados para su entrenamiento. Esta reparación se presume conducirá a reducir la variabilidad de *tokens* a utilizar, haciendo que, en primer lugar, el clasificador construido se pueda entrenar más rápidamente y, en segundo lugar, que los *tokens* que podrían generar "ruido" sean suprimidos.

Asimismo, se espera poder comparar las diferencias entre los *tokens* generados por esta implementación con aquellos generados por tokenizadores conocidos, como Gensim y NLTK. Esta comparación podrá realizarse no sólo sobre los textos de la fuente seleccionada, sino que podría realizarse sobre textos de otras fuentes, como por ejemplo Twitter, dado que hay estudios [20] que afirman que los resultados serían similares.

Así mismo, se buscan 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.

Formación de Recursos Humanos

En esta línea de I/D se han obtenido y se encuentran desarrollando actualmente dos becas de iniciación a la investigación. Asimismo, se esperan desarrollar una tesina de grado y una PPS, todas ellas dirigidas por miembros de este proyecto.

Bibliografía

[1] Tessore J. Esnaola L. Russo C. Ramón H. and Pompei S. 2018. Análisis automático de grandes volúmenes de datos en redes sociales mediante minería de textos combinado con algoritmos inteligentes. XX Workshop de Investigadores en Ciencias de la Computación. Universidad Nacional del Nordeste. Corrientes, Argentina.

[2] Xia Hu and Huan Liu. 2012. Mining Text Data. Text Analytics in Social Media. Springer, Boston, MA. DOI: https://doi.org/10.1007/978-1-4614-3223-4_12

[3] Charu C. Aggarwal and ChengXiang Zhai. 2012. Mining Text Data. An Introduction to Text Mining. Springer, Boston, MA. DOI: https://doi.org/10.1007/978-1-4614-3223-4_1

[4] Sarah Turnbull and Simon Jenkins.

2016. Why Facebook Reactions are good news for evaluating social media campaigns. Direct, Data and Digital Marketing Practice. (Feb. 2017), 17-156. DOI: https://doi.org/10.1057/dddmp.2015.56

[5] Henning Wachsmuth. 2015. Text Analysis Pipelines: Towards Ad-hoc Large-Scale Text Mining. Springer International Publishing. DOI: 10.1007/978-3-319-25741-9_2.

[6] Sophia Ananiadou and John Mcnaught. 2005. Text Mining for Biology And Biomedicine. Artech House Inc., Norwood, MA.

[7] Elizabeth Liddy. 2001. Encyclopedia of Library and Information Science, 2nd Ed. Natural Language Processing. Marcel Decker Inc., New York, NY.

[8] Eric S. Tellez, Sabino Miranda-Jimnez, Mario Graff, Daniela Moctezuma, Oscar S. Siordia, and Elio A. Villaseor. 2017. A case study of Spanish text transformations for twitter sentiment analysis. Expert Syst. Appl. 81, C (September 2017), 457-471. DOI: https://doi.org/10.1016/j.eswa.2017.03.071

[9] Kumar Ravi and Vadlamani Ravi. 2015. A survey on opinion mining and sentiment analysis. Know.-Based Syst. 89, C (November 2015), 14-46. DOI: https://doi.org/10.1016/j.knosys.2015.06.015

[10] Johan Bollen and Huina Mao. 2011. Twitter Mood as a Stock Market Predictor. Computer 44, 10 (October 2011), 91-94. DOI: http://dx.doi.org/10.1109/MC.2011.323

[11] Tetsuya Nasukawa and Jeonghee Yi. 2003. Sentiment analysis: capturing favorability using natural language processing. In Proceedings of the 2nd international conference on Knowledge capture (K-CAP '03). ACM, New York, NY, USA, 70-77. DOI: http://dx.doi.org/10.1145/945645.945658

[12] Alvaro Ortigosa, José M. Martín, and Rosa M. Carro. 2014. Sentiment analysis in Facebook and its application to e-learning. Comput. Hum. Behav. 31 (February 2014), 527-541.
DOI: http://dx.doi.org/10.1016/j.chb.2013.05.024

[13] Mifta Sintaha, Satter, S. Bin, Niamat Zawad, Chaity Swarnaker and Ahanaf Hassan. 2016. Cyberbullying detection using sentiment analysis in social media. Brac University, Dhaka, Bangladesh.

[14] Iñaki Alegria, Nora Aranberri, Pere R. Comas, Víctor Fresno, Pablo Gamallo, Lluis Padró, Iñaki San Vicente, Jordi Turmo, and Arkaitz Zubiaga. 2015. TweetNorm: a benchmark for lexical normalization of Spanish tweets. Lang. Resour. Eval. 49, 4 (December 2015), 883-905. DOI: http://dx.doi.org/10.1007/s10579-015-9315-6

[15] Bo Han, Paul Cook, and Timothy Baldwin. 2013. Lexical normalization for social media text. ACM Trans. Intell. Syst. Technol. 4, 1, Article 5 (February 2013), 27 pages.
DOI: http://dx.doi.org/10.1145/2414425.2414430

[16] Naradhipa, Aqsath Rasyid, and Ayu Purwarianti. 2012. Sentiment classification for Indonesian message in social media. International Conference on Cloud Computing and Social Networking. IEEE, 2012.

[17] Eleanor Clark and Kenji Araki. 2011. Text Normalization in Social them Media: Progress, Problems and Applications for a Pre-Processing System of Casual English. Procedia - Social and Behavioral Sciences, vol. 27. 2-11.

[18] Hunspell.github.io. 2019. Hunspell: About. Retrieved from: http://hunspell.github.io/.

[19] GNU Aspell. 2019. GNU Aspell. Retrieved from: http://aspell.net/.

[20] Timothy Baldwin, Paul Cook, Marco Lui, Andrew MacKinlay and Li Wang. 2013. How noisy social media text, how diffrnt social media sources?. Proceedings of the 6th International Joint Conference on Natural Language Processing (January 2013).

Análisis de tendencias en redes sociales

Aristides Dasso, Ana Funes

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}@unsl.edu.ar

RESUMEN

En este trabajo describimos algunos de los objetivos, lineamientos y métodos de una línea de investigación sobre la creación de modelos de opinión basados en datos provenientes de redes sociales, haciendo uso de técnicas de Minería de Datos tales como clasificación, agrupamiento, secuencias temporales y reglas de asociación.

Palabras clave: Redes sociales. Opinión en redes sociales. Métodos de Minería de Datos.

CONTEXTO

El trabajo de investigación aquí presentado se encuentra enmarcado dentro del ámbito del SEG (Software Engineering Group), de la Universidad Nacional de San Luis, ejecutándose dentro de una de las líneas de investigación del Proyecto de Incentivos código 22/F222 "Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad", dirigido por el Dr. Daniel Riesco y co-dirigido por el Dr. Roberto Uzal. El mismo se encuentra acreditado con evaluación externa y financiamiento de la Universidad Nacional de San Luis.

INTRODUCCIÓN

La extracción de conocimiento a partir de datos provenientes de redes sociales, tales como Facebook, Twiter, entre las más conocidas, con el objeto de determinar tendencias de opinión sobre temas de actualidad, es una práctica que se ha ido acentuando cada vez más en los últimos años y que parece haber venido para quedarse [10].

En este tipo de análisis de los datos, no solo se analizan opiniones sobre determinados temas de interés, sino que además se llega a analizar sentimientos [2], [3], [11]; se busca determinar líderes de opinión [8]; se analizan las preferencias de los votantes [6], [16] así como otros temas similares, todo esto además de su uso en investigaciones de mercado.

Por nuestra parte, en este trabajo de investigación, nos proponemos como objetivo realizar diversos análisis de opinión sobre temas de interés público, tales como temas económicos, políticos y otros como, por ejemplo, índice de precios, valor del dólar, etc., al poco tiempo que dichos datos aparezcan publicados en otros medios (TV, periódicos, etc.), para determinar su impacto en las redes sociales, analizando las opiniones vertidas por los usuarios en las mismas.

La intención es poder clasificar dichas opiniones, empleando técnicas de minería de datos y de textos, para así poder determinar tendencias y opiniones a lo largo de un cierto período de tiempo de cada una de los temas de interés analizados. Es decir, determinar cómo cambian las opiniones, como se

reciben, cada una de las noticias a lo largo del tiempo y explorar la posibilidad de establecer las razones de dichos cambios también a lo largo del tiempo.

LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Esta investigación sobre el análisis de tendencias, sentimientos, opiniones, empleando técnicas de la Minería de Datos y de Análisis de Texto (text mining), sobre datos obtenidos de redes sociales, se trata de una nueva línea de investigación dentro del marco del SEG (Sofware Engineering Group), que creemos reviste un gran interés y potencial de desarrollo, y en la cual nos encontramos dando los primeros pasos. Sin embargo, se trata de una extensión de una línea de investigación más amplia y consolidada dentro del grupo, que se ocupa de la aplicación y desarrollo de técnicas de Minería de Datos.

RESULTADOS Y OBJETIVOS

El objetivo principal en líneas generales que nos hemos propuesto es poder analizar, empleando técnicas de la Minería de Datos y de Análisis de Textos, las opiniones y sentimientos en las redes sociales frente a la aparición de noticias sobre sociedad, economía, política, etc. [1], [5], [12], [13], [14].

En este sentido, esperamos poder llevar a cabo diversos análisis que nos permitan identificar cuáles son las tendencias que se presentan en estas opiniones, si dichas opiniones y sentimientos se mantienen constantes en el tiempo, o si cambian, cómo cambian y en este caso, además, poder determinar las razones que llevan a dichos cambios.

Así, a modo de ejemplo, dado un tema de interés, como podría ser el aumento del IPC (Índice de Precios al Consumo), se pretende analizar de qué manera este tema impacta en las redes sociales a lo largo del tiempo; es decir, determinar cómo es recibido, y cómo el impacto o la opinión predominante sobre el mismo tema varía en el tiempo [7].

Asimismo, creemos que la extracción de conocimiento, plasmado en modelos inferidos a partir de datos recolectados en las redes sociales, que nos permitan establecer, bajo la forma de reglas de asociación, cuáles son las causas de dichos cambios, puede resultar también de interés.

FORMACIÓN DE RECURSOS HUMANOS

Dentro del SEG (Software Engineering Group), en el ámbito de la Universidad Nacional de San Luis, en el que se ejecuta el Proyecto de Incentivos código 22/F222 “Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad”, se vienen llevando a cabo numerosas tesis de grado y de posgrado.

En este sentido, creemos que la línea de investigación aquí descripta, la cual es una extensión de una línea más amplia sobre aplicación y desarrollo de técnicas de Minería de Datos, seguirá dando sus frutos, tanto en publicaciones nacionales e internacionales como en la formación de recursos humanos (2 tesis de maestría presentadas más una tesis de maestría en ejecución). Asimismo, de momento, se ha encarado la posibilidad de la ejecución de una nueva tesis de maestría basada en los objetivos que aquí nos hemos propuesto.

REFERENCIAS Y BIBLIOGRAFIA

[1] RAMSHREE, K.S. A Survey of Data Mining Techniques for Social Network Analysis. Software Engineering and Technology, [S.l.], v. 9, n. 1, p. 4-6, Feb. 2017. ISSN 0974 – 9632. http://www.ciitresearch.org/dl/index.php

[2] Abdalsamad Keramatfar, Hossein Amirkhani “Bibliometrics of sentiment analysis literature”. Journal of Information Science, vol. 45, 1: pp. 3-

15. , First Published March 19, 2018. DOI: 10.1177/0165551518761013

[3] B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis, Foundations and Trends" in Information Retrieval, vol 2, nos 1–2, pp 1–135, 2008. ISBN: 978-1-60198-150-9

[4] Chenhao Tan, Lillian Lee, Jie Tang, Long Jiang, Ming Zhou, Ping Li "User-Level Sentiment Analysis Incorporating Social Networks". Proceeding KDD '11 Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. Pages 1397-1405. San Diego, California, USA. August 21 - 24, 2011 ACM New York, NY, USA ©2011 ISBN: 978-1-4503-0813-7 doi 10.1145/2020408.2020614

[5] D. Kavitha, "SURVEY OF DATA MINING TECHNIQUES FOR SOCIAL NETWORKING WEBSITES". International Journal of Computer Science and Mobile Computing, Vol.6 Issue.4, April-2017, pg. 418-426

[6] Daniel José Silva Oliveira, Paulo Henrique de Souza Bermejo & Pâmela Aparecida dos Santos "Can social media reveal the preferences of voters? A comparison between sentiment analysis and traditional opinion polls," Journal of Information Technology & Politics, (2016) DOI: 10.1080/19331681.2016.1214094

[7] Daron Acemoglu, Giacomo Como, Fabio Fagnani, Asuman Ozdaglar. "Opinion fluctuations and disagreement in social networks". Journal Mathematics of Operations Research. Volume 38 Issue 1, February 2013, Pages 1-27 Institute for Operations Research and the Management Sciences (INFORMS), Linthicum, Maryland, USA

[8] Freimut Bodendorf, Carolin Kaiser. "Detecting Opinion Leaders and Trends in Online Social Networks", SWSM '09 Proceedings of the 2nd ACM workshop on Social web search and mining. Pages 65-68. Hong Kong, China, November 02 - 02, 2009. ACM New York, NY, USA ISBN: 978-1-60558-806-3 doi 10.1145/1651437.1651448

[9] J. A. Morente-Molinera, G. Kou, Y. Peng, C. Torres-Albero, E. Herrera-Viedma. "Analysing discussions in social networks using group decision making methods and sentiment analysis". author produced version of a paper published in: Information Sciences 447 (2018): 157-168. Repositorio Institucional de la Universidad Autónoma de Madrid. <https://repositorio.uam.es>

[10] Krzysztof Jędrzejewski, Mikołaj Morzy. "Opinion Mining and Social Networks: a Promising Match". International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2011, Kaohsiung, Taiwan, 25-27 July 2011. DOI: 10.1109/ASONAM.2011.123

[11] Lu Lin, Jianxin Li, Richong Zhang, Weiren Yu and Chenggen Sun. "Opinion Mining and Sentiment Analysis in Social Networks: A Retweeting Structure-aware Approach". 2014 IEEE/ACM 7th International Conference on Utility and Cloud Computing. 8-11 Dec. 2014. London, UK. Electronic ISBN: 978-1-4799-7881-6. DOI: 10.1109/UCC.2014.145 . Publisher: IEEE

[12] Mariam Adedoyin-Olowe, Mohamed Medhat Gaber and Frederic Stahl. "A Survey of Data Mining Techniques for Social Network Analysis." Journal of Data Mining & Digital Humanities (JDMDH), 2014

[13] Mohammad Noor Injadat, Fadi Salo1, Ali Bou Nassi. "Data Mining Techniques in Social Media: A Survey". Neurocomputing, June 2016.

[14] Norjihan Abdul Ghani, Suraya Hamid, Ibrahim Abaker Targio Hashem, Ejaz Ahmed. "Social media big data

analytics: A survey". Computers in Human Behavior (2018), https://doi.org/10.1016/j.chb.2018.08.039

[15] Pawel Sobkowicz, Michael Kaschesky, Guillaume Bouchard. "Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web". Government Information Quarterly 29(4):470–479. October 2012. DOI: 10.1016/j.giq.2012.06.005

[16] Pedro Saleiro; Sílvio Amir; Mário Silva; Carlos Soares. "POPmine: Tracking Political Opinion on the Web". 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing. Liverpool, UK. 26-28 Oct. 2015. DOI: 10.1109/CIT/IUCC/DASC/PICOM.2015.228

Aplicación de minería de datos e Internet de las Cosas (IoT) para investigaciones agropecuarias sobre artrópodos

Mariano R. Droz [(1)], Carlos E. Alvez [(1)], Graciela R. Etchart [(1)], Beatriz M. Diaz [(2)]

[(1)] Facultad de Ciencias de la Administración – Universidad Nacional de Entre Ríos
Av. Monseñor Tavella 1424 – (CP 3.200) Concordia, Entre Ríos, República Argentina

[(2)] Estación Experimental Agropecuaria Concordia – Instituto Nacional de Tecnología Agropecuaria
Ruta Provincial 22 y vías del Ferrocarril – (CP 3.200) Estación Yuquerí, Concordia, Entre Ríos, República Argentina

{drozm, caralv, getchart}@fcad.uner.edu.ar

diaz.beatriz@inta.gob.ar

RESUMEN

La Minería de Datos permite extraer información oculta y descubrir patrones y conocimientos interesantes a partir de grandes cantidades de datos. Las fuentes de los mismos pueden incluir bases de datos, *Data Warehouses*, la Web, u otros depósitos de información. Así, en el contexto actual, las tecnologías denominadas Internet de las cosas (IoT)[1] se han convertido en un proveedor de datos a gran escala. A partir de ello, se observa una estrecha relación entre la Minería de Datos e IoT, ya que las fortalezas de ambas permiten una complementación beneficiosa. Por ello, es posible encontrar nuevas aplicaciones en diferentes ámbitos. Algunas áreas donde dicha relación se está haciendo evidente, consiste en la agricultura de precisión y en la investigación agropecuaria. En esta última, existen oportunidades y nuevos campos de aplicación, como ser el estudio de la biodiversidad mediante la investigación sobre artrópodos en agroecosistemas. Resultando potencialmente factible aplicar Minería de Datos y utilizar una trampa de caída por tiempo con tecnología IoT, para determinar patrones de comportamiento de artrópodos, en función de factores abióticos relevantes para el campo de estudio.

Palabras clave: *Bases de Datos, Minería de Datos, Internet de las cosas (IoT), Agroecosistemas, Artrópodos.*

1 Estas tecnologías también son denominadas Internet de los Objetos o, en inglés, *Internet of Things*.

CONTEXTO

Este trabajo es desarrollado en el marco de una Tesis de la Maestría en Sistemas de Información de la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos. A su vez, está vinculado al Proyecto PID "Modelos de Machine Learning para la mejora de la precisión, seguridad y eficiencia en la gestión de datos biométricos" de la Universidad Nacional de Entre Ríos.

Por otra parte, se enmarca en la línea de trabajo "Desarrollo de tecnologías de bajo impacto ambiental aplicadas a la horticultura", que lleva a cabo el Grupo Hortícola de la Estación Experimental Agropecuaria Concordia del Instituto Nacional de Tecnología Agropecuaria.

1. INTRODUCCIÓN

La Minería de Datos (*Data Mining*) implica descubrir patrones novedosos, interesantes y potencialmente útiles de grandes conjuntos de datos y aplicar algoritmos para la extracción de información oculta. Muchas otras expresiones se usan para hacer referencia a las técnicas de Minería de Datos, entre ellas: descubrimiento de conocimiento en bases de datos (*Knowledge Discovery -mining- in Databases, KDD*), extracción de conocimiento (*knowledge extraction*), análisis de patrones (*data/pattern analysis*), arqueología de datos (*data archeology*) y recolección de información (*information harvesting*) [1], [2].

Una visión amplia de las técnicas de Minería de Datos permite definirlas como el proceso de descubrir patrones y conocimientos interesantes a partir de grandes cantidades de datos, donde las fuentes de los mismos pueden incluir bases de datos, *Data Warehouses*, la Web, u otros depósitos de información o datos [1].

A partir de esta última definición, actualmente es posible determinar que una fuente importante de datos para aplicar las técnicas de Minería de Datos, radica en las tecnologías denominadas Internet de las cosas (IoT).

El IoT puede concebirse como una infraestructura global de la sociedad de la información, que permite ofrecer servicios avanzados mediante la interconexión de objetos (físicos y virtuales), gracias a la interoperatividad de tecnologías de la información y la comunicación (TIC) presentes y futuras [3]. En el paradigma de IoT, muchos de los objetos que nos rodean están conectados de una forma u otra, generando enormes cantidades de datos que deben almacenarse, procesarse y presentarse de forma transparente, eficiente y fácilmente interpretable [4].

En este sentido, los datos recolectados por los dispositivos IoT se utilizan para comprender y controlar entornos complejos, lo que permite una mejor toma de decisiones, mayor automatización, eficiencia, productividad, precisión y reducción de costos [5]. Asimismo, es posible afirmar que este tipo de dispositivos se constituyen como una fuente de datos en permanente expansión en diferentes ámbitos de aplicación.

Por consiguiente, un desafío importante en este escenario, es el análisis oportuno de grandes cantidades de datos para producir conocimiento y decisiones altamente confiables y precisas, para que IoT pueda cumplir con sus expectativas [5].

Como consecuencia de lo anterior, surge la necesidad de hacer que IoT sea más inteligente, introduciendo tecnologías de análisis de datos. En efecto, una de las tecnologías más valiosas en este campo es la Minería de Datos [2], la cual ha tenido su origen a partir de la necesidad de descubrir automáticamente información valiosa de enormes cantidades de datos y transformar esos datos en conocimiento organizado [1].

De este modo, se observa una estrecha relación entre la Minería de Datos e IoT, ya que la aplicación de la misma en bases de datos recolectados o generados por dispositivos IoT permite y permitirá utilizar efectivamente ambas tecnologías en procura de objetivos comunes.

Las áreas de trabajo de estas tecnologías son vastas y cada vez es posible encontrar nuevas aplicaciones de las mismas en diferentes ámbitos. En lo que concierne a investigaciones agropecuarias, se evidencia un interés creciente en monitorear y recolectar datos sensibles del entorno, como ser: temperatura y humedad ambiente, temperatura y humedad del suelo, radiación ultravioleta (UV), nivel de precipitaciones, entre otros. Así, estos datos podrían ser correlacionados con otros de tipo temporal, de posicionamiento y de rendimiento de cultivos, con el fin de descubrir patrones ocultos, optimizar, predecir y aportar precisión a determinados procesos productivos, área denominada agricultura de precisión.

Investigaciones recientes han obtenido resultados promisorios [6], [7], desarrollando sistemas de riego inteligentes, basados en una red de sensores IoT que recolectan datos relevantes para el problema, aplicándose técnicas de Minería de Datos con el propósito de generar nuevo conocimiento, mejorar la toma de decisiones y el proceso productivo mediante una optimización de la gestión del riego. Otro aporte interesante, consiste en el desarrollo de un prototipo de un sistema de control para el cultivo hidropónico del tomate [8], el cual utiliza tecnología IoT y desarrolla e implementa una *Deep Neural Network* para controlar con mayor precisión el sistema en tiempo real.

Cabe señalar también, que otra línea específica de investigación agropecuaria consiste en el estudio de la biodiversidad. En este campo, alcanzar una mayor comprensión sobre la dinámica de las especies y sobre la estructura y funcionamiento de los ecosistemas resulta ser apremiante, dada la incertidumbre sobre la dimensión de los posibles efectos del cambio climático sobre la misma [9], [10].

En esta línea de investigación, un área de trabajo comprende el estudio sobre artrópodos en agroecosistemas, donde la idea básica subyacente al uso de artrópodos como indicadores de características generales de los ecosistemas es que, la sola presencia (o abundancia) de algunas especies es el producto de una larga y compleja lista de circunstancias que la han hecho posible [11]. Así, la existencia o no de determinadas especies de artrópodos da cuenta de la biodiversidad de un agroecosistema específico.

En este campo de estudio, es habitual la utilización de trampas de caída (*pitfall trap*) y de trampas de caída por tiempo (*time-sorting pitfall trap*). El avance tecnológico también ha permitido la evolución de estas trampas [12], [13], mediante la incorporación de sensores y unidades de registro locales de datos (*datalogger*). Específicamente [13], ha desarrollado una trampa que registra la fecha, la hora, el número del tubo muestreado y la temperatura del suelo. Posteriormente, se relacionan estos datos con la identificación y conteo de los artrópodos capturados en distintos experimentos. A su vez, en su trabajo también se menciona la posibilidad de expandir este tipo de trampas mediante la incorporación de sensores adicionales.

De este modo, surge el interés por desarrollar un prototipo que incorpore nuevos sensores para relevar factores abióticos relevantes para el campo de estudio. Sumado a lo anterior, la posibilidad de dotar de tecnología IoT abre nuevas posibilidades a dichos dispositivos, permitiendo complementar metodologías de investigación específicas de la disciplina, descubriendo y determinando patrones de actividad de los artrópodos del suelo, mediante la aplicación de Minería de Datos en una base de datos específicamente diseñada para este fin.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente trabajo se enmarca en el análisis y aplicación de técnicas de Minería de Datos con fines científicos. Asimismo, cuenta con un enfoque transversal, ya que explorará la relación entre estas técnicas e IoT, con el propósito de diseñar una base de datos que almacenará datos recolectados por una trampa de caída por tiempo dotada de tecnología IoT, la cual registrará factores abióticos tales como temperatura y humedad ambiente, temperatura y humedad del suelo y radiación ultravioleta (UV).

A partir de ello, específicamente el trabajo apunta a la aplicación de Minería de Datos e IoT para el estudio de la biodiversidad e impacto del cambio climático, mediante la investigación agropecuaria sobre patrones de comportamiento de artrópodos del suelo.

Por otra parte, esta propuesta se encuadra en la línea de investigación *"Bases de Datos",* establecida como una de las líneas prioritarias dentro de la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos, según las Resoluciones "C.D." Nº 160/11 y 203/11.

Asimismo, se relaciona con las denominadas AgroTIC, según lo expresado en el *"Libro Blanco de la Prospectiva TIC - Proyecto 2020"* [14]. Concretamente, involucra distintas áreas de trabajo, entre ellas: bases de datos y gestión del conocimiento, y sistemas de adquisición y administración de datos incorporados a dispositivos de uso específico.

3. RESULTADOS ESPERADOS

El objetivo general consiste en describir, analizar y aplicar Minería de Datos en una base de datos específicamente diseñada para dispositivos científicos con tecnología IoT utilizables en investigaciones agropecuarias sobre patrones de comportamiento de artrópodos.

Para cumplir con dicho objetivo, se presentarán y clasificarán las principales técnicas de la Minería de Datos. A su vez, se describirán y analizarán las características relevantes del IoT. Luego de ello, se profundizará en la aplicación de estas técnicas a partir de datos obtenidos por dispositivos con tecnología IoT.

A continuación, se creará el diseño conceptual de una trama de caída por tiempo

con nuevos sensores y tecnología IoT apta para investigaciones agropecuarias sobre artrópodos. Posteriormente, se diseñará una base de datos para recolectar los datos generados por dicho dispositivo. Así, la misma será puesta a prueba mediante simulaciones de captura y almacenamiento de datos, las cuales darán paso a una etapa de aplicación de las técnicas de Minería de Datos con el propósito de efectuar una determinación de patrones de comportamiento de artrópodos del suelo en función de factores abióticos.

De este modo, el aporte original del trabajo consiste en aplicar una innovación al utilizar la Minería de Datos para la determinación de patrones de comportamiento de artrópodos del suelo, a partir de datos recolectados por una trampa de caída por tiempo con nuevos sensores y tecnología IoT, respecto a lo cual aún no se observa evidencia científica donde se hayan utilizado ambas tecnologías para dicha finalidad.

4. FORMACIÓN DE RECURSOS HUMANOS

Las tareas se desarrollarán como parte de las actividades planificadas para que uno de los integrantes realice la Tesis de la Maestría en Sistemas de Información de la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos, acreditada por la Resolución CONEAU Nº 1276/12.

A su vez, proporcionará un marco propicio para la formación e iniciación en actividades de investigación, mediante la realización de Proyectos de Investigación Científica, Desarrollo e Innovación Tecnológica en el marco del Programa de Investigación y Desarrollo la Universidad Nacional de Entre Ríos.

De igual modo, será un ámbito pertinente para que estudiantes de la Universidad Nacional de Entre Ríos, de carreras de grado afines, lleven a cabo sus proyectos de Trabajo Final o participen como Becarios de Iniciación en la Investigación.

5. BIBLIOGRAFÍA.

[1] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2011.

[2] A. V. Vasilakos, J. Wan, F. Chen, X. Rong, D. Zhang, and P. Deng, “Data Mining for the Internet of Things: Literature Review and Challenges,” *Int. J. Distrib. Sens. Networks*, vol. 11, no. 8, p. 431047, 2015.

[3] UIT-T, “UIT-T Rec. Y.2060 Descripción general de Internet de los objetos,” *Sect. Norm. las Telecomunicaciones la UIT*, no. 2012-06-15, p. 20, 2012.

[4] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Futur. Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, 2013.

[5] F. Alam, R. Mehmood, I. Katib, and A. Albeshri, “Analysis of Eight Data Mining Algorithms for Smarter Internet of Things (IoT),” *Procedia Comput. Sci.*, vol. 58, no. DaMIS 2016, pp. 437–442, 2016.

[6] A. Goap, D. Sharma, A. K. Shukla, and C. Rama Krishna, “An IoT based smart irrigation management system using Machine learning and open source technologies,” *Comput. Electron. Agric.*, vol. 155, no. September, pp. 41–49, 2018.

[7] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, and P. Nillaor, “IoT and agriculture data analysis for smart farm,” *Comput. Electron. Agric.*, vol. 156, no. November 2018, pp. 467–474, 2019.

[8] M. Mehra, S. Saxena, S. Sankaranarayanan, R. J. Tom, and M. Veeramanikandan, “IoT based

hydroponics system using Deep Neural Networks," *Comput. Electron. Agric.*, vol. 155, no. October, pp. 473–486, 2018.

[9] E. Uribe, "Estudios del cambio climático en América Latina: El cambio climático y sus efectos en la biodiversidad en América Latina," *Euroclima*, vol. vol 51, no. 3, p. 86, 2015.

[10] D. U. Hooper, E. C. Adair, B. J. Cardinale, J. E. K. Byrnes, B. A. Hungate, K. L. Matulich, A. Gonzalez, J. E. Duffy, L. Gamfeldt, and M. I. O'Connor, "A global synthesis reveals biodiversity loss as a major driver of ecosystem change," *Nature*, vol. 486, p. 105, May 2012.

[11] I. Ribera and G. Foster, "El uso de artrópodos como indicadores biológicos," *Bol. la Soc. Entomológica Aragon.*, vol. 20, pp. 265–276, 1997.

[12] S. Buchholz, "Design of a time-sorting pitfall trap for surface-active arthropods," *Entomol. Exp. Appl.*, vol. 133, no. 1, pp. 100–103, 2009.

[13] M. S. McMunn, "A time-sorting pitfall trap and temperature datalogger for the sampling of surface-active arthropods," *HardwareX*, vol. 1, pp. 38–45, 2017.

[14] G. Baum, A. Artopoulos, C. Aguerre, I. Albornoz, V. Robert, and MINCYT, *Libro Blanco de la prospectiva TIC, Proyecto 2020*. 2009.

Aplicaciones Informáticas para el Estudio de Biodiversidad de Poliquetos Espiónidos en los Golfos Nordpatagónicos

Gustavo Samec[1,2,3], María Emilia Diez[1,2,5], Marcos Zárate[1,2,4], Carlos Buckle[1,2], Joaquín Lima[1,2], Rodrigo Jaramillo[1,2,3], Alejandro Sánchez[1,2,6], Renato Mazzanti[1,2,3]

[1] Departamento de Informática, Facultad de Ingeniería, Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB), Puerto Madryn.
[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) (CCT CONICET-CENPAT)
[5] Laboratorio de Parasitología (LAPA), Instituto de Biología de Organismos Marinos, Centro Nacional Patagónico, Consejo Nacional de Investigaciones Científicas y Técnicas (IBIOMAR) (CCT CONICET-CENPAT)
[6] Departamento de Informática, Universidad Nacional de San Luis (UNSL), San Luis.
{gsamec, emiliadiez, zarate, renato}@cenpat-conicet.gob.ar, cbuckle@unpata.edu.ar

RESUMEN

En la actualidad, la información biológica es de tal magnitud que se hace necesario el uso de herramientas informáticas para optimizar la labor del científico. En este sentido, una base de datos regional constituye una buena herramienta para la validación y la precisión de datos con estándares de control. En las costas patagónicas uno de los principales problemas para la marisquería es el causado por los poliquetos espiónidos polidorinos. Una correcta identificación de estas especies es clave para atender los problemas de infestación. El objetivo general es estudiar teorías y técnicas que permitan el desarrollo de herramientas que asistan a biólogos en la validación de las especies involucradas en sus estudios. Se propone realizar un proceso de ingeniería de software para atender el caso particular de espiónidos de los golfos nordpatagónicos mediante la incorporación de información de distribución, herramientas avanzadas de consulta y semántica para la interoperabilidad con repositorios de especies de referencia mundial que garanticen la calidad de los datos. El proyecto está integrado por especialistas de Ciencias de la Computación y Biológicas, conformando un grupo de investigación interdisciplinario en informática para la biodiversidad dentro del Laboratorio de Investigación en Informática (LINVI) de la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB).

Palabras clave: Base de Datos, Taxonomía, Poliquetos, Spionidae, Web Semántica.

CONTEXTO

Este proyecto fue realizado por especialistas de Ciencias de la Computación y de las Ciencias Biológicas, que conforman un grupo de investigación interdisciplinario en informática para la biodiversidad dentro del LINVI de la UNPSJB. En los últimos años Orensanz J.M., Mazzanti R., Samec G. y Lima J. han desarrollado la base de datos Southwest Atlantic Benthic Invertebrates[1] (SWATL) con la finalidad de centralizar la información sobre invertebrados bentónicos de la región con sus datos asociados para

[1] http://sistema.cenpat-conicet.gob.ar:8081/sgrb/sgrb/system/Menu.jsf

hacer interpretaciones y realizar aplicaciones prácticas en el campo de la conservación. Sobre esta base se pretende el desarrollo de herramientas informáticas que faciliten el estudio de la taxonomía y ecología de estas especies. Para poder llevar a cabo el proyecto será necesario especificar, diseñar y desarrollar un sistema de bases de datos que considere la semántica y la interoperabilidad con repositorios de referencia mundial de especies, como por ejemplo Global Biodiversity Information Facility[2] (GBIF) y World Register of Marine Species[3] (WoRMS). En este caso en particular se trabajará con uno de los grupos más abundantes y diversos de la comunidad bentónica como es la clase Polychaeta. Este proyecto apunta a consolidar un sistema de información de registros biológicos como un repositorio de registros sobre especies de poliquetos espiónidos para la región patagónica, con características de base de datos abierta e interoperable, que por su implementación genérica pueda incorporar a futuro otras especies de otros grupos.

1. INTRODUCCIÓN

Existen iniciativas a nivel global como GBIF, Especies2000[4] o WoRMS que gestionan listas taxonómicas de especies cuyo objetivo es la gestión de índices uniformes y la validación por especialistas taxónomos. Como antecedente directamente relacionado con el proyecto se cuenta con la base de datos NONATObase[5] que trabaja con registros de especies de clase Polychaeta del Atlántico Sudoccidental (y áreas cercanas al antártico), desde 5°N a 80°S y cuenta con un comité científico externo de asesoramiento. La fuente de datos primarios que utilizan se basa en la exploración y recopilación de información desde distintos sitios Web como Web of *Knowledge*[6], Scopus[7], Scielo[8] y desde librerías y/o repositorios de distintas instituciones de investigación. Dentro del contexto nacional el Ministerio de Ciencia, Tecnología e Innovación Productiva (MinCyT) creó el Sistema Nacional de Datos Biológicos[9] (SNDB) sobre una plataforma desarrollada por GBIF basada en el estándar *Darwin Core* [1]. Además creó el Sistema Nacional de Datos del Mar (SNDM) con el fin de visualizar la información de los centros nacionales productores de datos marinos de Argentina de tipo satelitales, geológicos, batimétricos, físico-químicos, geofísicos y biológicos. Sumado a esto en mayo de 2014 se lanza oficialmente la iniciativa Pampa Azul[10], y se propone utilizar la plataforma del SNDM para el manejo del conocimiento generado en los puntos focales de la iniciativa (Golfo San Jorge, Islas Sándwich del Sur, Agujero Azul y Talud, Banco Burwood/Namuncurá). La gestión de la biodiversidad requiere el uso de información biológica heterogénea de múltiples fuentes. La indexación, la agregación y la búsqueda de dicha información se basan en los nombres y el conocimiento taxonómico de los organismos. Sin embargo, las taxonomías cambian con el tiempo debido a nuevos hallazgos científicos. Además, los nombres de las especies y su significado cambian en el tiempo, diferentes autoridades usan diferentes nombres científicos para un mismo taxón en diferentes momentos, y varios nombres vernáculos están en uso en diferentes idiomas.

[2] https://www.gbif.org/
[3] http://www.marinespecies.org/
[4] https://sp2000.org/
[5] http://nonatobase.ufsc.br/
[6] https://www.webofknowledge.com/
[7] https://www.scopus.com
[8] http://www.scielo.org
[9] http://datos.sndb.mincyt.gob.ar/
[10] http://www.pampazul.gob.ar/

Esto dificulta la integración de datos y la recuperación de información si no se dispone de información biológica detallada. Entre los trabajos más importantes podemos mencionar [2,3,4]. De acuerdo con [5] el surgimiento de la Web Semántica [6] 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. En lo que refiere a los datos científicos, los datos abiertos enlazados [7] se presentan como el camino estándar hacia la integración de datos heterogéneos de diferentes dominios. Este proyecto apunta a desarrollar un sistema de información que contenga datos de diversidad de poliquetos (Annelida: Polychaeta) de la familia Spionidae sobre la base de datos SWATL. La finalidad de este sistema es aportar herramientas para la gestión taxonómica y de la información asociada a cada registro.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El presente proyecto plantea el desarrollo de herramientas que asistan a biólogos y otros actores interesados a la validación de las especies involucradas en sus estudios de poliquetos, mediante la taxonomía y propone especificar y diseñar un sistema de bases de datos, con el agregado de tecnologías semánticas y estándares de *World Wide Web Consortium* (W3C) como RDF[11], SPARQL[12], etc. y el estándar GeoSPARQL[13] del *Open Geospatial Consortium* (OGC) para la interoperabilidad con registros de referencia mundial de especies.

Por otro lado, el proyecto de investigación *Infraestructura de Acceso a Datos Primarios con aporte de semántica en Repositorios Digitales* [8] tiene varios objetivos generales. Uno de ellos es el de desarrollar conocimiento especializado en el área de Interoperabilidad Semántica de la Información. En este sentido, se estudian, entre otras, el desarrollo de agentes de información cuyo ambiente de trabajo es la Web, además de técnicas de representación de conocimiento, razonamiento automático y modelado ontológico.

Estas líneas de investigación confluyen en el estudio de formalismos y tecnologías para cubrir las necesidades emergentes de compartir, actualizar e integrar el conocimiento de sistemas computacionales pre-existentes como lo es SWATL.

3. RESULTADOS ESPERADOS

En este proyecto se estudiarán teorías y se desarrollarán modelos y aplicaciones de software para asistir a biólogos que trabajan en biodiversidad, contribuyendo a la validación de las especies involucradas y aportando calidad a sus registros. En este caso particular se trabajará con un grupo de organismos muy abundante y diverso de la comunidad bentónica como son los poliquetos espiónidos de los golfos nordpatagónicos.

Objetivos particulares:

1. Analizar el estado del arte respecto de la interoperabilidad con repositorios o bases de datos de especies de referencia mundial.
2. Analizar y rediseñar el modelo de datos de la base de datos SWATL para la implementación de un sistema de información.

[11] https://www.w3.org/RDF/
[12] https://www.w3.org/TR/sparql11-query/
[13] https://www.opengeospatial.org/standards

3. Facilitar mediante herramientas informáticas la obtención de sinonimias, las fuentes y referencias para un taxón, como así también la distribución geográfica y sus atributos.
4. Estudiar y aplicar tecnologías de la Web Semántica para la publicación de datos abiertos enlazados, interoperabilidad semántica y explotación de información.
5. Consolidar un grupo de investigación interdisciplinario en informática para la biodiversidad dentro del grupo LINVI de la UNPSJB.

En este sentido, el proyecto permitirá el desarrollo de un sistema de información taxonómica/biológica, que constituirá una herramienta para facilitar y mejorar la calidad de los estudios taxonómicos y ecológicos, en un continuo flujo de retroalimentación de la información [9].

La relevancia de la propuesta se identifica a diferentes niveles:

1. Tecnológico: La herramienta a desarrollar utiliza tecnologías semánticas para la integración con diferentes repositorios internacionales. Esto será de gran relevancia no sólo para investigadores taxónomos o ecólogos, sino también para aquellos organismos estatales o instituciones que trabajan en temáticas de conservación de la biodiversidad, de manejo de recursos pesqueros entre los principales.
2. Socioeconómico: La infestación de polidorinos afecta a moluscos de interés comercial en los golfos nordpatagónicos. Esto es relevante y pertinente en tanto es necesario contar con información sobre el estado de salud de los moluscos como satisfacer estándares de calidad del producto.
3. Política científica: Contar con un sistema que gestione este tipo de información facilitará la investigación de métodos combinados de control biológico y prácticas de manejo para una explotación más eficiente.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan integrantes de formación docente y académica en las áreas de Ingeniería de Software, Bases de Datos, Inteligencia Artificial y Biología. Cinco de los docentes son del Departamento de Informática y uno del Departamento de Matemáticas de la Facultad de Ingeniería de la UNPSJB Sede Puerto Madryn. La Doctora en biología M. Emilia Diez es especialista en taxonomía y ecología de poliquetos, dos de los autores están realizando carreras de doctorado y otros dos se encuentran desarrollando carreras de especialización y maestrías. Uno de los autores de este proyecto 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 y otro está finalizando su doctorado en Ciencias de la Computación en la Universidad Nacional de San Luis. También forman parte del proyecto un graduado de la carrera de Licenciatura en Informática y tres 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, este proyecto tiene la intención de guiarlos en el desarrollo de sus tesinas en esta rama de la disciplina. Otro aporte para la formación académica radica en la posibilidad de que los alumnos puedan realizar Instancias Supervisadas de Formación en la Práctica Profesional en el marco de este proyecto y en pasantías y/o terrenos de capacitación en el área de Biología. Si bien el proyecto de investigación es nuevo, continúa con las líneas de investigación iniciadas en otros proyectos anteriores. El objetivo es combinar los esfuerzos realizados en forma individual en un proyecto común.

5. BIBLIOGRAFÍA

1. Wieczorek J, Bloom D, Guralnick R, Blum S, Döring M, Giovanni R, et al. Darwin Core: an evolving community-developed biodiversity data standard. PLoS One. 2012;7: e29715.

2. Tuominen J, Laurenne N, Hyvönen E. Biological Names and Taxonomies on the Semantic Web – Managing the Change in Scientific Conception [Internet]. The Semanic Web: Research and Applications. 2011. pp. 255–269. doi:10.1007/978-3-642-21064-8_18

3. Michel F, Faron-Zucker C, Tercerie S, Olivier G. TAXREF-LD: A Reference Thesaurus for Biodiversity on the Web of Linked Data [Internet]. Proceedings of TDWG. 2017. p. e20232. doi:10.3897/tdwgproceedings.1.20232

4. Plitzner P, Henning T, Müller A, Güntsch A, Karam N, Kilian N. Bottom-up taxon characterisations with shared knowledge: describing specimens in a semantic context. S4BioDiv@ ISWC. 2017. Available: https://pdfs.semanticscholar.org/7f02/4d5917825e07370ac3f19df4ae4429fdf98a.pdf

5. Koper R. Use of the Semantic Web to Solve Some Basic Problems in Education: Increase flexible, distributed lifelong learning; decrease teacher's workload [Internet]. Journal of Interactive Media in Education. 2004. p. 5. doi:10.5334/2004-6-koper

6. Berners-Lee T, Hendler J, Lassila O, Others. The semantic web. Sci Am. New York, NY, USA:; 2001;284: 28–37.

7. Heath T, Bizer C. Linked Data: Evolving the Web into a Global Data Space. Synthesis Lectures on the Semantic Web: Theory and Technology. Morgan & Claypool Publishers; 2011;1: 1–136.

8. Mazzanti R, Zàrate M, Samec G, Buckle C. Infraestructura de acceso a datos primarios con aporte de semántica en repositorios digitales. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). 2018. Available: http://sedici.unlp.edu.ar/handle/10915/67259?show=full

9. Escalante T, Llorente J, Espinosa D, Soberón J. Bases de datos y sistemas de información: aplicaciones en biogeografía. Rev Acad Colombiana Cienc Exact Fis Natur. 2000;24: 325–341.

Aplicando métodos y técnicas de la ciencia de los datos a datos universitarios

Herrera, M.[1,2], Ruiz, S.[1], Romagnano, M.[1,2], Ganga, L.[1], Lund, M.[1,2] y Torres, E.[1]

[1]Dpto. de Informática / FCEFN/ Universidad Nacional de San Juan/ {myriamhrrr, leonelganga, etorres271}@gmail.com, sbruiz@yahoo.com.ar

[2]Instituto de Informática/ FCEFN/ Universidad Nacional de San Juan/ {mlund, maritaroma}@iinfo.unsj.edu.ar

RESUMEN

En la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFN) de la Universidad Nacional de San juan (UNSJ), específicamente en cada departamento que agrupa las distintas carreras universitarias, es de gran interés conocer las características sociales y culturales de los estudiantes universitarios que permitan determinar un perfil socio-económico del estudiante por departamento y a su vez brindar conocimiento a la Universidad de los hábitos de vida y estudio de ellos, permitiendo a la institución ajustar sus líneas de acción y decisiones en función de las necesidades de los estudiantes. Este conocimiento incide positivamente en la mejora de las condiciones y calidad de vida de la comunidad educativa desde su ingreso, permanencia en la carrera, hasta su egreso.

En pos del conocimiento de los factores que influyen tanto en la calidad de vida de los estudiantes como así también aquellos que repercuten en la institución en forma directa como ser rendimiento académico, grado de satisfacción del egresado, se han utilizado métodos y técnicas de la Ciencia de los Datos.

***Palabras clave*:** análisis de datos, ingresantes-alumnos-egresados, métodos y técnicas de clasificación.

CONTEXTO

El presente trabajo se encuentra enmarcado en el proyecto denominado "Calidad Educativa Universitaria Mediante Técnicas de Clasificación".

Este proyecto fue aprobado por evaluadores externos, en la Convocatoria 2017 de la UNSJ, para el período comprendido entre 01/01/2018 al 31/12/2019. Se desarrolla en el Gabinete de Matemática Aplicada del Instituto de Informática de la FCEFN de la UNSJ.

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". Vigencia: 01/01/2016 -31/12/2017. Código: 21 E/1011.
- "Algoritmos de Clasificación de Procesos Multivariados utilizando Medidas de Asociación Espacial". 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". Vigencia: 01/01/2014 – 31/12/2015. Código 21 F/982.
- "Clasificación Espacial Multivariada". Vigencia: 01/01/2011 – 31/12/2013. Código: 21 E/ 878
- "Reducción y Selección de Variables en la Clasificación Digital". Vigencia: 01/01/2008 - 31/12/2010. Código: 21 E/820.
- "Aplicación de una metodología en la medición de la calidad del proceso

enseñanza aprendizaje en la universidad". Vigencia: 01/05/2003 al 31/12/2005.

1. INTRODUCCIÓN

Identificar la tipología del alumnado universitario es indispensable para determinar la calidad universitaria, permitiendo sensibilizar y orientar ciertas acciones. Este concepto ha ido asumiendo protagonismo en diferentes esferas. Específicamente, en el ámbito de la educación se relaciona con la propuesta de mejorar la organización educativa atendiendo a lineamientos que contemplen el desarrollo de modelos en los que la currícula académica sea funcional y contextualizada [3, 15].

En pos del conocimiento de los factores que influyen tanto en la calidad de vida de los estudiantes como así también aquellos que repercuten en la institución en forma directa como ser rendimiento académico y grado de satisfacción del egresado, se han utilizado técnicas y métodos propios de la Ciencia de los Datos [4, 8, 14].

Así mismo, para descifrar patrones y tendencias en los datos se utiliza tanto el análisis de conglomerados como el análisis discriminante, técnicas que algunos autores ubican entre las más potentes para aplicar en investigaciones de índole social. Este último tipo de método permite clasificar sujetos u objetos a partir de características similares. Los mismos se diferencian en la manera en la cual extraen conocimiento útil a partir de los datos [5].

El Análisis Discriminante (AD) cuenta con grupos de datos conocidos, así como observaciones de unidades cuya pertenencia a los grupos, en términos de los grupos conocidos, es desconocida inicialmente y tiene que ser determinada a través del análisis de los datos. Este tipo de problemas de clasificación es referido como reconocimiento de patrones asistido, en terminología estadística cae bajo el título de Análisis Discriminante. Por otro lado, hay problemas de clasificación donde los grupos son desconocidos a priori y el principal propósito del análisis es determinar los grupos a partir de los propios datos, de modo que las unidades, dentro del mismo grupo, sean en algún sentido más similares u homogéneas que aquellas que pertenecen a grupos diferentes. Dicho problema de clasificación es referido como reconocimiento de patrón no supervisado y en terminología estadística cae bajo el título de Análisis de Conglomerados (AC) [9].

Además, el Análisis de Datos Simbólicos (ADS) y de Lógica Difusa (LD) son también herramientas para minar datos, los cuales generalizan los métodos clásicos exploratorios e informáticos, ya que en muchas actividades humanas es común recopilar grandes volúmenes de datos en grandes bases [11, 12, 13].

Hoy en día, la disponibilidad de grandes volúmenes de información y el uso generalizado de herramientas informáticas, han transformado el análisis de datos orientándolo hacia determinadas técnicas especializadas de Minería de Datos o Data Mining (DM). Las técnicas de MD persiguen el descubrimiento automático del conocimiento contenido en la información almacenada de modo ordenado en grandes bases de datos. Tienen como objetivo de detectar patrones de comportamiento consistentes o relaciones entre las diferentes variables para aplicarlos a nuevos conjuntos de datos [2, 13]. Para realizar este proceso no trivial de identificación válida, novedosa, potencialmente útil y entendible de patrones comprensibles que se encuentran ocultos en los datos se utilizan técnicas estadísticas que permiten reunir y analizar los datos, como también técnicas provenientes de la inteligencia artificial. Tanto las técnicas estadísticas como las técnicas de inteligencia artificial son bastante poderosas. En algunos casos, estas son tan solo dos enfoques o alternativas diferentes para la solución de un mismo problema, en otras ocasiones son técnicas complementarias porque resuelven problemas de diferente naturaleza [6, 10].

Así, en la MD se encuentran estrategias de descubrimiento de reglas de asociación, que tienen como base la lógica clásica, y se proponen encontrar conjuntos de elementos

que co-ocurren juntos frecuentemente en una base de datos [1,7].
En el actual proyecto se están aplicando las técnicas mencionadas para analizar algunos factores influyentes en la calidad universitaria, como así también se espera detectar tipologías básicas de grupos, obtenidos a partir de ingresantes, alumnos y egresados de la FCEFN de la UNSJ.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Cuando se habla de minar los datos se hace referencia al conjunto de métodos, técnicas y tecnologías que permiten explorar grandes bases de datos, de manera automática o semiautomática, con el objetivo de encontrar patrones repetitivos, tendencias o reglas que expliquen el comportamiento de los datos en un determinado contexto.
El Data Mining surge para intentar ayudar a comprender el contenido de un repositorio de datos. Con tal fin, fundamentalmente hace uso de prácticas estadísticas y, en algunos casos, de algoritmos de búsqueda; próximos a la Inteligencia Artificial y a las Redes Neuronales.
Por lo tanto, y fundado en los párrafos precedentes, el grupo de investigación se encuentra trabajando en las siguientes líneas de investigación:

- minería de datos enfocada en aspectos estadísticos, sistemas de computación y lógica difusa,
- identificación y análisis de tendencia de acuerdo a datos universitarios,
- análisis de aplicación del estudio anterior a otras áreas de aplicación; por ejemplo la financiera,
- estudio y análisis del funcionamiento del software SODAS a fin de lograr emular su funcionamiento mediante el desarrollo de una aplicación de uso libre, usando como lenguaje de programación a Phyton y/o R,
- modelado de información cualitativa de forma aproximada, a través de la lógica difusa, y
- evaluación de la calidad universitaria; mediante la definición de un instrumento de medición.

3. OBJETIVOS ESPERADOS Y RESULTADOS OBTENIDOS

Como objetivo general se espera analizar los datos de egresados que provienen del sistema on line Siu Kolla y de los alumnos, los que provienen de la encuesta procesada.

Como objetivos específicos se han propuesto:

- Formar recursos humanos, al nivel de grado y posgrado, en la temática que involucra la calidad de la Institución.
- Determinar las variables influyentes (poder de discriminación y predicción) que caractericen al ingresante, a los alumnos y a los egresados universitarios.
- Transferencia de herramienta metodológica a las instituciones educativas del medio.

Conjuntamente, hasta el momento, se han logrado las siguientes divulgaciones y/o publicaciones:

- "Tipología Influyente en el Rendimiento Académico de Alumnos Universitarios", Autoras: S. Ruiz, M. Herrera, M. Romagnano, L. Mallea y M. I. Lund. Revista Entre Ciencia e Ingeniería, año 12 (vol 12), Nº 23 – Primer semestre de 2018, páginas 109–116. ISSN 1909-8367 (Impreso), ISSN 2539-4169 (En línea). DOI: http://dx.doi.org/10.31908/19098367.3710.
- "Calidad Universitaria mediante Técnicas del Data Mining". XX Workshop de Investigadores en Ciencias de la Computación – XX WICC 2018 (UNNE,

Corrientes, 26 y 27 de Abril). Autores: Myriam Herrera, María Inés Lund, Susana Beatriz Ruiz, María Gema Romagnano. ISBN 978-987-3619-27-4.http://wicc2018.unne.edu.ar/wicc2018librodeactas.pdf

- "Rendimiento académico aplicando técnicas del análisis multivariado". VI Jornadas Nacionales y II Latinoamericanas de Ingreso y Permanencia en carreras Científico-Tecnológicas", IPECyT 2018. 16, 17 y 18 de mayo de 2018. Olavarría – Buenos Aires, Autores: Ruiz, Susana; Herrera, Myriam; Romagnano, María; Lund, María Inés. Libro de resúmenes: https://drive.google.com/file/d/1jPCSYWy6z-Y4_Esvc1AhE4yxhhUWdj6l/view.
- "Análisis Simbólico de Datos: una potente herramienta para Big Data". VI Jornadas de Cloud Computing & Big Data, JCC&BD 2018. 25-29 de Junio. La Plata. Autores: Adriana Mallea, Myriam Herrera y María Inés Lund. http://jcc.info.unlp.edu.ar/index.php/resumenes-aceptados/
- "Aprendizaje Colaborativo en un curso de Ingeniería de Software: Medición de la Satisfacción". II Congreso Internacional de Ciencias de la Computación y Sistemas de Información 2018 – CICCSI 2018. San Juan, Mendoza del 5 al 9 de Noviembre de 2018. Autores: Laura N. Aballay y María Inés Lund. Proceedings en elaboración.
- "Calidad Universitaria: Tipología y Variables Influyentes". Autores: Myriam Herrera, María Romagnano, Leonel Ganga y Susana Ruiz. VI Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación (SABTIC 2018), Universidad Tecnológica Nacional (UTN), Facultad Regional de Concepción del Uruguay (Entre Ríos, Argentina), 25 y 26 de octubre de 2018. Disponible en: http://www.frcu.utn.edu.ar/sabtic/es/.
- "Clasificación de clientes de tarjeta de crédito de una institución financiera a partir de WEKA". Autores: Leonel Ganga, Myrian Herrera y María Romagnano. Congreso Internacional de Ciencias de la Computación y Sistemas de Información (CICCSI2018) – Universidad de Campagnat, Mendoza y Universidad Nacional de San Juan, San Juan, Argentina, 5-6 y 8-9 de noviembre de 2018.

4. FORMACIÓN DE RECURSOS HUMANOS

En el grupo de trabajo se encuentran dos doctorandos, un magister, un alumno de la Licenciatura en Ciencias de la Computación en etapa de culminación de escritura de tesis y un ayudante-alumno de la carrera de Licenciatura en Ciencias de la Computación. Se prevé la incorporación de alumnos de grado y posgrado por medio de trabajos finales.

5. BIBLIOGRAFÍA

[1] Agrawal, R., Imielinski, T. y Swami, A. "Mining association rules between sets of items in large databases". In: Proceedings of the ACM SIGMOD International Conference on Management of Data. Washington D.C., 1993, p. 207-216.

[2] Anaya, A. y Boticario, J. G. "A data mining approach to reveal representative collaboration indicators in open collaboration frameworks", 2009.

[3] Di Gresia, L., Fazio, M., Porto, A., Ripani L., y Escudero, W. S. "Rendimiento y productividad de los estudiantes. El caso de las universidades públicas

argentinas," Departamento de Economía, Facultad de Ciencias Económicas, Universidad Nacional de La Plata, La Plata, 2004.

[4] Garbanzo Vargas, G. M. "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.

[5] González López, I. "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.

[6] Gurmendi, M. de L., "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.

[7] Hahsler, M., Hornik, Kurt; Reutterer, Thomas (2006). Implications of probabilistic data modeling for mining association rules. En From Data and Information Analysis to Knowledge Engineering. Springer Berlin Heidelberg. p. 598-605.

[8] Hipp, J., Güntzer, U.; Nakhaeizadeh, G. "Algorithms for association rule mining—a general survey and comparison. ACM sigkdd explorations newsletter", vol. 2, no 1, p. 58-64, 2002.

[9] Johnson, R. A. y Wichern, D. W. "Applied multivariate statistical analysis". Pearson Prentice Hall, 2007.

[10] Marín Llanes L. A., Carro Cartaya J. C. "La Minería de Datos como Herramienta en el Proceso de Inteligencia Competitiva" Consultoría Biomundi, Dirección de Inteligencia Corporativa, Instituto de Información Científica y Tecnológica (IDICT), CUBA, Taller Nacional sobre Inteligencia Empresarial IntEmpres'2000 [Online] available: http://www.redciencia.cu/empres/index.htm

[11] Moscoloni, N. "Las Nubes de Datos. Métodos para analizar la complejidad," 2005.

[12] Peña, D. "Análisis de datos multivariantes". McGraw-Hill/Interamericana, 2002.

[13] R. A. (Richard A. Johnson and D. W. Wichern. "Applied multivariatestatistical analysis". Pearson Prentice Hall, 2007.

[14] Rojas Diaz, J., Chavarro Porras, J., Moreno Laverde, R. "Técnicas de lógica difusa aplicadas a la minería de datos. Scientia et technica, [S.l.], v. 3, n. 40, ISSN 2344-7214, 2009. Disponible en: <http://revistas.utp.edu.co/index.php/revistaciencia/article/view/3095/1695>. Fecha de acceso: 05 nov. 2018 doi:http://dx.doi.org/10.22517/23447214.3095.

[15] Torrado-Fonseca, M. M. y Berlanga-Silvente, V. "Revista d'innovació i recerca en educació" Rev. d'Innovació i Recer. en Educ., vol. 6, no. 2, pp. 150–166, 2013.

Aportes a Nuevos Modelos de Bases de Datos

Jorge Arroyuelo, Maria E. Di Genaro, Alejandro Grosso, Verónica Ludueña, Nora Reyes
Dpto. de Informática, Fac. de Cs. Físico-Matemáticas y Naturales, Universidad Nacional de San Luis
{*bjarroyu, digeme, agrosso, vlud, nreyes*}@unsl.edu.ar

Edgar Chávez
Centro de Investigación Científica y de Educación Superior de Ensenada, México
elchavez@cicese.mx

Rodrigo Paredes
Dpto. de Cs. de la Computación, Fac. de Ingeniería, Universidad de Talca, Chile
raparede@utalca.cl

Resumen

La evolución de las tecnologías de información y comunicación, junto con la gran cantidad y variedad de información disponible digitalmente han llevado al surgimiento de nuevos depósitos no estructurados de información, para los datos que no se adaptan fácilmente al modelo relacional. Estos datos surgen desde campos muy disímiles provocando requerimientos de usuarios que pueden ser tan dispares como el tipo de datos que se necesita procesar. Por tal motivo, es necesario utilizar estos depósitos especializados y formas más sofisticadas de búsqueda sobre los mismos, ya que las soluciones tradicionales no suelen enfrentar tales requerimientos. Para satisfacer estas demandas se deben desarrollar aplicaciones capaces de manipular eficientemente datos muy diferentes entre sí como: secuencias biológicas, huellas digitales, texto, audio, video, imágenes, etc.

Por otro lado, la gran cantidad de datos que se debe 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 *conscientes de la jerarquía de memoria*. Un modelo que se adapta a tales requerimientos, en el cual se puede utilizar métodos de acceso que contemplen estos aspectos, son las *Bases de Datos Métricas*. Esta investigación pretende contribuir a la madurez de este nuevo modelo de bases de datos considerando distintas perspectivas.

Palabras Claves: bases de datos no convencionales, índices, búsquedas por proximidad.

Contexto

El presente trabajo se realizó en el marco del Proyecto Consolidado *Tecnologías Avanzadas de Bases de Datos*, (Cód. 03-2218 y en Programa de Incentivos 22-F814) de la Universidad Nacional de San Luis, en la línea *Bases de Datos no Convencionales*. En colaboración con investigadores del Centro de Investigación Científica y de Educación Superior de Ensenada (México) y de la Universidad de Talca (Chile).

La investigación que se realiza en este ámbito, está enfocada en lograr la consolidación de las bases de datos destinadas a manipular datos no convencionales. Esto incluye además, plantear nuevas arquitecturas del procesador que mejoren a muy bajo nivel los administradores de estos nuevos modelos de 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.

Introducción

La inserción de la computación en casi todos los ámbitos de la sociedad: productivo, artístico, laboral, recreativo, científico, de la salud, etc., hizo necesaria la evolución de las bases de datos. éstas deben 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.

El modelo de *espacios métricos* resulta adecuado para englobar muchas de las 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 entre sí son los objetos. Este contexto permite abordar aplicaciones tales como el ingresar un trozo de melodía en un buscador y que éste presente melodías similares a dicho trozo, o proveerle una imagen y esperar que éste proporcione imágenes semejantes a la suministrada. Las búsquedas tradicionales (exactas) carecen de sentido en la mayoría de estos casos y sobre estos tipos de datos, resultando más naturales las *búsquedas por similitud*, provistas por este modelo.

Para responder eficientemente a este tipo de búsquedas, sin realizar la examinación exhaustiva del conjunto de datos, se necesitan los llamados *Métodos de Acceso Métricos* (MAMs). Sin embargo, además de poder responder eficientemente las búsquedas, es esencial su actualización y optimización ya que la mayoría de estos métodos no admiten actualizaciones (inserciones/eliminaciones), ni están diseñados para soportar conjuntos masivos de datos y tampoco para resolver operaciones de búsquedas más complejas. Los avances en este campo se ven reflejados en áreas tales 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.

Además, en otra de las áreas exploradas, se pretende 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 a bajo nivel de los administradores de bases de datos (DBMS).

Líneas de Investigación y Desarrollo

Bases de Datos no Convencionales

En este ámbito de investigación, las bases de datos que administran videos, imágenes, texto libre, secuencias de ADN o de proteínas, audio, etc., llamadas *bases de datos no convencionales*, serán modelizadas utilizando el modelo de espacios métricos. Además, a fin de responder eficientemente consultas por similitud se hará uso de MAMs. Debido a lo costoso que generalmente resultan los cálculos de distancia, el número de cálculos realizados al crear el índice o al realizar búsquedas es considerado 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 más comunes son de dos tipos: por *rango* o de k*-vecinos más cercanos* (k-NN).

Métodos de Acceso Métricos

Muchas veces los índices no caben en memoria principal, ya sea porque organizan una base de datos masiva, o porque los objetos de la misma son demasiado grandes. Entonces, surge la necesidad de diseñar índices especialmente pensados para memoria secundaria. Teniendo esto en consideración, se han diseñado dos nuevos índices basados en la *Lista de Clusters*(*LC*) [5] que son totalmente dinámicos, es decir, admiten inserciones y eliminaciones de elementos y están especialmente diseñados para trabajar sobre grandes volúmenes de datos [7]. 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 para la *DSC*, que está en etapa de evaluación, 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 dichos 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 e 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, por consiguiente, reduciendo los tiempos de acceso. Esta nueva variante, denominada *BOLDSC* (por sus sigla en inglés "Bu?ered On Line Dynamic Set of Clusters"), se espera guarde las buenas características de dinamismo de *DSC*, mejorando su desempeño tanto en construcción como en búsquedas.

Entre la optimizaciones necesarias para los MAM's también se encuentra el dinamismo. Por esta razón, a partir de uno de los índices de mejor desempeño en espacios de mediana a alta dimensión, el *árbol de Aproximación Espacial* (*SAT*), totalmente estático, se desarrolló el *árbol de Aproximación Espacial Dinámico (DSAT)* [7] que permite realizar inserciones y eliminaciones, conservando un muy buen desempeño en las búsquedas, pero agregando un parámetro a sintonizar. Una variante también estática del *SAT*, el *árbol de Aproximación Espacial Distal* (*DiSAT*) [4], logra optimizar las búsquedas respecto de ambos (*SAT* y *DSAT*) y además no necesita ningún parámetro. Por ello, se ha propuesto una optimización del mismo, la *Foresta de Aproximación Espacial Distal* (*DiSAF*) [2], 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, como en este método cada inserción de un elemento provoca la reconstrucción parcial de la estructura, es costosa de construir. Por ello, se ha diseñado una nueva versión dinámica del *DiSAT*, que demora las reconstrucciones para amortizar su costo entre muchas inserciones. A este nueva forma de inserción se la denomina *inserción perezosa*. La ventaja es que logra bajar los costos de construcción, manteniendo el buen desempeño en las búsquedas.

Otra faceta que hay que tener en cuenta, son los requerimientos de algunas aplicaciones que priorizan la rapidez en las respuestas aunque sea a costa de perder algunos elementos: se intercambia precisión (devolviendo sólo algunos objetos relevantes) por velocidad 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. En este contexto, se están analizando distintas posibilidades de utilizar búsquedas aproximadas para acelerar las búsquedas, pero intentando mantener una buena calidad de respuesta.

Problema de los k Vecinos

Entre las búsquedas por similitud más utilizadas, la de los k vecinos más cercanos resulta útil en la predicción de funciones, en el aprendizaje automático, la cuantificación y compresión de imágenes, entre otras. Una generalización de la misma consiste en obtener los k-vecinos más cercanos de *todos* los elementos de la base de datos (*All*-k-NN). El planteo general de esta consulta puede verse como: para cada elemento $u \in X$, donde la base de datos $X \subseteq \mathbb{U}$, se responde con los k elementos en X-$\{u\}$ que tengan la menor distancia d a u. La solución burda a este problema tiene una complejidad de n^2 cálculos de distancia, con $|X| = n$; ésta consiste en comparar cada elemento de la base de datos con todos los demás. Una solución más eficiente implica el preprocesamiento de los datos, por ejemplo por medio de un índice, para reducir el número de cálculos de distancia en las búsquedas.

El *Grafo de los k-vecinos más cercanos* (kNNG) [8] se encuentra entre las soluciones propuestas para espacios métricos generales y su desempeño supera algunas de las técnicas clásicas. La construcción del kNNG permite indexar un espacio métrico y luego él mismo se emplea en la resolución de las consultas por similitud. Sin embargo, existen situaciones en las cuales el costo de la construcción del índice, para luego obtener los vecinos más cercanos, puede resultar excesivo. Por ejemplo, cuando la función de distancia es demasiado costosa de calcular, o si se tiene una base de datos masiva, o cuando se pretende 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. Otro factor a tener en cuenta son los requerimientos de algunas aplicaciones particulares, que priorizan la velocidad de respuesta sobre la precisión de la misma [9, 5, 10, 6]. Para hacer frente a éstas circunstancias es que se han considerado las llamadas *búsquedas por similitud aproximadas*. Este tipo de consultas aceptan algunos "errores" en la respuesta, si con esto se mejora la complejidad de la misma.

Para tal fin, se han desarrollado algunas propuestas que computan una aproximación del kNNG, es decir conectan cada objeto u de la base de datos con k vecinos *cercanos*, relajando la condición que exige que no haya, en toda la base de datos, algún objeto más cercano a u que los k vecinos devueltos. Esto puede ocasionar que se pierda algún objeto muy cercano y en su lugar se devuelva otro un poco más lejano, pero a cambio la respuesta será más rápida. A este grafo se lo denominó *Grafo de vecinos cercanos* (knNG) [3]. Una característica común que tienen estos desarrollos es que ninguno utiliza un índice para buscar en él. Una primera aproximación aprovecha el profundo conocimiento que se tiene del *DiSAT* para plantear un enfoque novedoso. Aquí se consideró un caso particular del problema, cuando $k = 1$, obteniendo el 1nNG. Esta propuesta utiliza la información obtenida durante la *construcción* del *DiSAT* para construir el 1nNG, conectando a cada elemento u con un elemento cercano de la base de datos, que puede ser, o no, su vecino más cercano [3]. 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.*

Las otras propuestas abordadas se enfocan en responder a los *All-k*-nN y computar el knNG. Estos planteos no utilizan el apoyo de ningún índice, no sólo no buscan en ellos, sino que ni siquiera recurren a la información provista por su construcción. El propósito de estos desarrollos es aprovechar de manera ingeniosa las propiedades de la *función de distancia*. En ellos se sugieren distintas maneras de seleccionar muestras de la base de datos, a partir de las cuales se obtiene un conjunto de distancias que serán el punto de partida de este proceso. También se analizan diferentes maneras de utilizar la información conseguida, para calcular los vecinos aproximados para todos los objetos de la base de datos, utilizando propiedades como la simetría o la desigualdad triangular. Los resultados de estas propuesta se muestran muy prometedores.

Arquitecturas de Procesadores Orientadas a Bases de Datos

Hay autores que realizan una distinción entre arquitectura, implementación y realización. La arquitectura de una computadora define el conjunto mínimo de propiedades que determinan qué programas correrán y qué resultados producirán sobre el procesador, es decir es la interfaz entre el software y el hardware de una computadora. La organización básica del flujo de datos y el control, conforma la implementación. La estructura física que comprende la implementación, conforma la realización [1].

En la actualidad, la investigación sobre arquitecturas de procesadores ha sido desplazada por investigación sobre la implementación de procesadores. La mayoría de los trabajos de investigación están dedicados a mejorar técnicas de predicción (tanto de control como de datos), técnicas para sincronizar y comunicar procesadores (núcleos) a través de 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. No obstante, estas técnicas de implementación se pueden aplicar a todo tipo de arquitecturas, 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.

Si bien las arquitecturas RISC compitieron en desempeño con las arquitecturas CISC [2], las mismas poseen un alto tráfico de bits entre el procesador y la memoria para una determinada traza de ejecución. Esto finalmente favoreció a las CISC sobre las RISC, una vez que las CISC mejoraron sus técnicas de implementación. El objetivo de las investigaciones en esta línea de trabajo es plantear nuevas arquitecturas que minimicen el tráfico de bits entre el procesador y la memoria. Para esto se está realizando la construcción de un simulador del set de instrucción AMD-64 o x86-64 con el fín de evaluar el tráfico de bits para benchmarks como Specint y Specfp para la arquitectura x86. El próximo paso sería evaluar el tráfico de bits para la arquitectura propuesta sobre los mismos benchmarcks, lo cual implica construir no sólo el simulador de la arquitectura sino también el compilador C para la misma. Finalmente, se pretende aprovechar el conocimiento adquirido para, desde bajo nivel, mejorar el desempeño de los DBMSs.

Resultados y Objetivos

Los resultados obtenidos en las investigaciones realizadas sobre el modelo de espacios métricos, además de permitir mejorar el desempeño de los MAMs analizados, conducen a estudiar su aplicación a otros métodos de acceso [2, 3, 4, 7].

[1] acrónimo del inglés “Reduced Instruction Set Computer”.
[2] Acrónimo del inglés “Complex Instruction Set Computer”.

Se ahondará en el estudio de nuevos diseños de estructuras de datos, capaces de adaptarlas tanto al nivel de la jerarquía de memorias donde se almacenarán, como a las características de los datos a ser indexados, mejorando así su eficiencia. Se espera brindar nuevas herramientas de administración para bases de datos métricas eficientes, que permitan una madurez cada vez más cercana a la de los modelos tradicionales de base de datos.

Se profundizará el análisis sobre cómo resolver consultas de manera más eficiente, sin la utilización de índices, y sobre distintos espacios. Además, se espera mejorar el desempeño de las operaciones de bajo nivel en los DBMS, mediante una nueva arquitectura del procesador.

Actividades de Formación

Dentro de esta línea de investigación se forman alumnos y docentes-investigadores en:

- **Doctorado en Cs. de la Computación:** una tesis en desarrollo sobre expresividad de lenguajes lógicos de consulta.
- **Maestría en Cs. de la Computación:** una tesis sobre búsqueda por similitud aproximada, próxima a su defensa, y otra en desarrollo sobre un índice dinámico eficiente.
- **Maestría en Informática**: una tesis finalizada, de la Universidad Nacional de San Juan, sobre un índice dinámico para búsquedas aproximadas, especialmente diseñado para memoria secundaria.

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, M. Di Genaro, N. Reyes, and P. Roggero. Decomposability of disat for index dynamization. *Computer Science & Technology*, pages 110–116, 2017.

[3] E. Chávez, V. Ludueña, N. Reyes, and F. Kasián. All near neighbor graph without searching. *Computer Science & Technology*, 18(1):61–67, April 2018.

[4] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal {SAT}. *Information Systems*, 59:15 – 47, 2016.

[5] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM Computing Surveys*, 33(3):273–321, September 2001.

[6] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.

[7] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.

[8] 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.

[9] 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.

[10] P. Zezula, G. Amato, V. Dohnal, and M. Batko. *Similarity Search: The Metric Space Approach (Advances in Database Systems)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2005. XVIII, 220 p., Hardcover ISBN: 0-387-29146-6.

Aportes vinculados a Bases de Datos Espaciales y Espacio Temporales

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

La línea de investigación *Bases de Datos Espaciales y Espacio Temporales* del proyecto *Tecnologías Avanzadas de Bases de Datos,* está orientada a vincular las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas. El objetivo general es utilizar técnicas, métodos y herramientas para investigación de base, así como el desarrollo de aplicaciones para contribuir en la resolución de problemas provenientes de diferentes dominios de aplicación.

Palabras clave: Bases de Datos, Geometría Computacional, Metaheurísticas, Bases de Datos Espaciales y Espacio Temporales, AgroTIC.

CONTEXTO

El proyecto Tecnologías Avanzadas de Bases de Datos desarrolla actividades relacionadas al tratamiento de objetos, estructurados y no estructurados, utilizados en diversos campos de aplicación, tales como sistemas de información geográfica, computación gráfica, computación móvil, robótica, diseño asistido por computadora, motores de búsqueda en internet, entre otras.

En el citado 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 no tradicionales 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.

En particular, los modelos de bases de datos espaciales y espacio temporales se utilizan en aplicaciones en las cuales se necesita guardar y consultar información actual e histórica de posiciones referenciadas espacialmente y cambios de forma que tuvieron los objetos de estudio en diferentes escenarios a lo largo del tiempo. De esta manera, se plantea la necesidad de la incorporación de métodos científicos, procesos y sistemas de descubrimiento de información oculta en grandes cantidades de datos estructurados y no estructurados, para obtener información a los efectos de contribuir en la toma de decisiones en un sistema.

En este contexto, se vinculan las disciplinas Bases de Datos, Geometría Computacional y Metaheurísticas con la finalidad de recurrir a métodos y herramientas que aporten en la resolución de problemas de diversos dominios

de aplicación y para la resolución de problemas orientados a optimización.

El trabajo de investigación de la línea se desarrolla en estrecha colaboración con investigadores afines de proyectos de la Universidad Nacional de San Luis, convenios entre organizaciones nacionales y provinciales con presencia en San Luis, como así también de universidades extranjeras mediante convenios de cooperación interinstitucional.

1. INTRODUCCIÓN

Las Tecnologías de la Información y la Comunicación (TIC) generan mejoras en diversas actividades humanas y continúan incorporando cambios sociales, económicos y culturales, entre otros, convirtiéndose en procesos clave para accionar sobre el presente y proyectar hacia el futuro. Así, diferentes situaciones del mundo real ameritan el uso de las mismas, en pos de una mejora sustancial en cuestiones de calidad, gestión, economía, etc.

En particular, con el objetivo de estudiar la potencialidad y alcance real de las TIC en el ámbito de la producción agropecuaria se integró el proyecto *Campo Conectado*, que surgió a partir de la inquietud de varias instituciones, entre ellas el Instituto Nacional de Tecnología Agropecuaria (INTA-Estación Experimental Agropecuaria San Luis) y la Universidad Nacional de San Luis [1].

Otro eje de trabajo, que define objetivos de investigación está relacionado con problemas de optimización combinatoria, para los cuales no se puede garantizar encontrar la mejor solución en un tiempo razonable, para todas las instancias del problema. La optimización procura encontrar la mejor solución posible a un problema de este tipo dentro de un período de tiempo limitado. 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 [2].

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 [10].

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 [8, 11]. La Triangulación de Dilación Mínima (Minimum Dilation Triangulation, MDT) donde la dilación mide la calidad de conexión entre puntos de la triangulación es otro problema estudiado. Para este problema no se conoce un algoritmo que lo resuelva en tiempo polinomial y tampoco se ha demostrado que sea NP-duro [9].

La utilización de 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.

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.
- Aplicación de la Geometría Computacional y su marco disciplinar, para considerar aspectos propios de los problemas involucrados.
- Optimización de estructuras geométricas relacionadas con las bases de datos citadas, mediante la aplicación de metaheurísticas para la optimización de problemas NP-duros en Geometría Computacional.

- Desarrollo de herramientas para la visualización de información y aplicaciones vinculadas con bases de datos espacio temporales.

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.
- 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 vinculadas a las bases de datos mencionadas.

3. RESULTADOS OBTENIDOS

En el marco del proyecto 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.

En este contexto, se presentó una propuesta de desarrollo de una plataforma integral, soporte para diversos eventos y sistemas de información con dominio de aplicación en el sector agropecuario, mediante el uso de las TIC. Sus características principales están orientadas a la recolección de datos de diferentes fuentes con almacenamiento compartido, integración progresiva de diversas funcionalidades, explotación y visualización de la información [5].

En el marco de esta plataforma integral se trabaja 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. La herramienta realiza la comunicación con las bases de datos y las consultas espaciales y espacio temporales de los objetos de tratamiento, así como la visualización del geoseguimiento disponible por medio de tecnología móvil [7].

Conjuntamente se propone el desarrollo de una herramienta de apoyo en la gestión de rodeos de cría, que permite la recolección de datos desde diferentes fuentes para la inferencia y el seguimiento de la condición corporal de los individuos del rodeo. La herramienta provee funcionalidades para el análisis de datos, generación de indicadores, obtención de conocimiento de apoyo para la toma decisiones y visualización de la información disponible por medio de tecnología móvil [6].

Estas herramientas, actualmente en etapa de desarrollo, estarán incluidas en la plataforma integral para permitir realizar un seguimiento del rodeo y generar indicadores que aporten información relevante para el proceso de toma de decisiones. Dicha plataforma se propone disponible en la web, con accesibilidad mediante tecnología móvil (I+D+i).

En relación al desarrollo de herramientas para la visualización vinculadas a las bases de datos mencionadas, se propuso como objetivo la difusión del material fotográfico perteneciente a José La Vía alojado en la Fototeca de la Universidad Nacional de San Luis. Dicho objetivo se logró con la creación de una aplicación que permite enlazar el pasado con el presente, permitiendo a los usuarios redescubrir la ciudad de San Luis desde una óptica distinta. Este desarrollo se encuentra orientado a dispositivos móviles con sistema operativo Android y haciendo hincapié principalmente en el uso de Realidad Aumentada para lograr una experiencia atractiva al transitar por la ciudad, poniendo en valor al fondo fotográfico custodiado por la Fototeca. El desarrollo de estas herramientas

se plasmó en un trabajo final de Licenciatura en Ciencias de la Computación.

Con respecto a problemas de optimización abordados con técnicas metaheurísticas, se retomó el problema MDT para el cual, actualmente, se está utilizando el modelo de Programación Evolutiva introducido por L. J. Fogel [4], con el fin de desarrollar nuevos operadores y obtener resultados que se puedan comparar con los obtenidos anteriormente. Este modelo se basa en el comportamiento adaptativo de las nuevas generaciones más que en la modificación de cromosomas de cada individuo. La Programación Evolutiva utiliza sólo dos operadores: mutación y selección, es decir que no se realiza recombinación. Los resultados serán contrastados con los obtenidos con los algoritmos: Greedy, Local Search (LS), Iterated Local Search (ILS), Simulated Annealing (SA), y Random Local Search (RLS) presentados en [3].

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo en la Universidad Nacional de San Luis, consolida su formación 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 reflejada en tesis doctorales, tesis de maestría y Licenciados en Ciencias de la Computación.

ii) Actividades de formación académica, a través del dictado y realización de cursos de posgrado y de especialización.

iii) Realización de jornadas de investigación con docentes de otras universidades.

iv) Actividades de divulgación científica, conferencias y publicaciones en congresos.

v) Integración del proyecto interinstitucional Campo Conectado.

La línea tiene como un objetivo continuar con las actividades relacionadas al presente proyecto, proponiendo nuevas actividades de formación académica, de formación de recursos humanos, investigación, desarrollo, y otras actividades académico-científicas vinculantes.

5. BIBLIOGRAFÍA

[1] *Campo Conectado*, proyecto interinstitucional. http://inta.gob.ar/noticias/campo-conectado-un-nuevo-proyecto-interinstitucional (2017).

[2] de Berg, M., Cheong O., van Kreveld, M., Overmars, M.,. *Computational Geometry: Algorithms and Applications.*, Springer-Verlag, Heidelberg (2008).

[3] Dorzán M. G. et al. *Approximated algorithms for the minimum dilation triangulation problem*. Journal Heuristics 189–209 (2014).

[4] Fogel, L. J. *Autonomous automata.* Industrial Research 4, 14-19 (1962).

[5] Gagliardi, E., Dorzán, M. G., Taranilla, M. T., Palmero, P., Casanova, C.: *Propuesta de plataforma para la integración de TIC orientadas al Agro* Anales de CAI 2017 Congreso de AgroInformática, 46JAIIO (2017)

[6] Gagliardi, E..; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *Diseño de una herramienta de apoyo en la gestión de rodeos de cría.* Anales de CAI 2018 Congreso de AgroInformática, 47JAIIO (2018).

[7] Gagliardi, E.; Dorzán, M.G.; Taranilla, M.T.; Palmero, P.; Casanova, C.: *GeoSeguimiento de Rodeos, hacia una plataforma integral para el Agro*. Anales de CAI 2018 Congreso de AgroInformática, 47JAIIO (2018).

[8] Gudmundsson J., Levcopoulos C.; *Minimum weight pseudo-triangulations*. Computational Geometry. Theory and applications. Elsevier Vol. 38- Pages 139-153 (2007).

[9] Knauer, C., Mulzer, W.: *Minimum Dilation Tiangulations.* Tech. rep., FreieUniversitt Berlin, Fachbereich Mathematik und Informatik (2005).

[10] Michalewicz Z., Fogel D., *How to Solve It: Modern Heuristics*, Springer, (2004).

[11] Mulzer W., Rote G. *Minimum weight triangulation is NP-hard.*. Proceedings of the 22 Annual ACM Symposium on Computational Geometry (2006).

AVANCES EN EL DESARROLLO DE MÉTODOS DE DESAMBIGUACIÓN Y RECOMENDACIÓN DE AUTORES CIENTÍFICOS PARA UN METABUSCADOR DE LAS CIENCIAS DE LA COMPUTACIÓN

H. Kuna, A. Cantero, A. Canteros, M. Rey, E. Zamudio, 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

En el ámbito de la recuperación y procesamiento de datos relacionados a publicaciones científico-tecnológicas existen diversos problemas a resolver. En este trabajo son de especial interés aquellos relacionados con el tratamiento de datos de los autores de tales publicaciones.

La identificación unívoca de un autor, la constitución de un perfil del mismo compilando su historial de publicaciones, un conjunto de métricas que permitan estimar su impacto en la comunidad científica y la selección de un subconjunto de autores a partir de una consulta del usuario de un metabuscador, son algunos de los problemas que se tratan en la presente línea de investigación.

En este trabajo se presentan los avances en dos soluciones planteadas: las correspondientes a un método de desambiguación de autores y, por otra parte, un método de recomendación de autores basado en clases previamente definidas.

Mediante los resultados que se reseñan en este documento se pretende contribuir a la mejora del desempeño de procesos de explotación de información asociados a la recuperación de producciones científico-tecnológicas.

Palabras clave: producción científico-tecnológica, autores científicos, 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 información relativa a producciones científicas a través de internet es uno de los desafíos actuales de la actividad científica. Los grandes volúmenes de publicaciones, la diversidad de fuentes de consulta y de herramientas de ayuda al investigador junto a la expansión de los datos que se generan a partir de la interacción entre científicos, son las bases de este problema.

Entre las diferentes herramientas que ayudan al investigador a recuperar datos de calidad y relevantes se pueden mencionar: los buscadores científicos, los repositorios digitales, redes sociales del ambiente y soluciones de procesamiento de grandes bases de datos [1], [2]. Dentro de este último grupo, se agrupan herramientas que proveen algunos de los siguientes servicios [1], [3]: integración de resultados; generación automática de perfiles de autores, publicaciones o centros de investigación; visualización de redes de colaboración entre autores; clasificación automática de contenidos; cálculo de métricas relativas al impacto o reconocimiento de autores y/o publicaciones; comparación de perfiles de entidades, entre otros.

En la presente línea de investigación se han abordado algunos de los problemas mencionados y se ha desarrollado un Sistema de Recuperación de Información (SRI),

concretamente un metabuscador, que opera sobre documentos científicos del área de Ciencias de la Computación [4]–[6]. En el desarrollo del SRI se han planteado diferentes avances en torno a la conformación de una base de datos (BD) interna para el metabuscador que permitiera la generación de soluciones como las mencionadas en el apartado anterior [7]–[9]. En el desarrollo de esta BD se ha identificado uno de los ejes del presente trabajo, el problema de la desambiguación de entidades [9] y haciendo uso del contenido de la misma se ha generado el método de recomendación de autores que completa los avances a reseñar.

Se denomina entidades a todos los actores involucrados en los procesos del metabuscador, de los cuales se han definido perfiles para el almacenamiento de sus datos y su disponibilidad para implementar procesos que hicieran uso de los mismos para contribuir a mejorar los resultados que el metabuscador presenta a sus usuarios [7]. La gestión de este tipo de contenido requiere implementar procesos que ayuden a identificar unívocamente aquellas entidades asociadas a las producciones científicas. En paralelo, se consideró adecuado asistir al investigador generando recomendaciones de entidades que pudieran resultar relevantes para su investigación. En función de este objetivo se generó el método de recomendación de autores que considera diferentes aspectos de los mismos para establecer una clasificación [8] y posteriormente elabora un listado para presentar al usuario aquellos que pueden ser de utilidad dada su consulta.

1.1 DESAMBIGUACIÓN DE AUTORES

La identificación de un autor de una publicación científica es primordial para conocer la obra del mismo y poder obtener referencias a trabajos previos que permitan conocer su historial de trabajo. En la actualidad esta actividad no puede ser desarrollada manualmente por los grandes volúmenes de datos disponibles, es así como surgen diferentes procesos y métodos que tienen por objeto facilitar esta tarea. Tales recursos son los que se agrupan dentro de lo que se denomina desambiguación de entidades. En una instancia inicial, dentro del presente grupo de investigación se ha determinado comenzar por la desambiguación de los autores registrados en la BD del metabuscador. Casos como: publicaciones de un mismo autor con distintos nombres, diferentes autores con un mismo nombre, ausencia de datos en los registros de la publicación, entre otros, son los que generan inconsistencias en la identificación unívoca de autores.

En la literatura del área de ciencias de la computación se encuentran diferentes métodos que constituyen alternativas de solución para este problema. Entre las técnicas de base utilizadas se pueden identificar métodos de agrupamiento y de clasificación. En todos los casos se plantea la necesidad de contar con un volumen adecuado de datos para asegurar la calidad de los resultados [10]–[12].

1.2 RECOMENDACIÓN DE AUTORES MEDIANTE PERFILES

Las tareas de búsqueda de expertos en un área de conocimiento son conocidas en la comunidad científica. La construcción de perfiles de expertos o *expert profiling* busca tomar y organizar datos a fin de evidenciar la experiencia de los autores en una temática en particular. Estas evidencias podrán variar en términos de las fuentes de datos empleadas para la construcción del perfil, sin embargo, las publicaciones científico-tecnológicas de un autor se consideran el mejor reflejo de su desempeño.

Para contar con una BD de autores se comenzó por el diseño de los perfiles a través de los que serían registrados los mismos y sus publicaciones. Posteriormente se desarrollaron de métodos de extracción, transformación y carga (ETL, por su sigla en inglés) para obtener datos desde diferentes fuentes y unificarlos en la BD generada. Sobre este conjunto de datos se generó una serie de clases en las cuales se etiquetó a los autores en función de sus antecedentes, para luego determinar su recomendación al usuario [8]. En este trabajo se presentan los avances relacionados a las acciones finales del método correspondientes con la selección de los autores a recomendar.

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.

Actualmente se está trabajando sobre dos tareas derivadas de tal objetivo. Los procesos de desambiguación de entidades, en particular de autores, son necesarios para obtener datos de calidad sobre los cuales luego se pueda aplicar algún tipo de técnica de reconocimiento de patrones. Por otra parte, la recomendación automática de autores constituye un primer acercamiento a los objetivos planteados, permitiendo presentar al usuario del metabuscador datos relativos a su consulta que van más allá de las publicaciones que vayan a ser recuperadas.

El aspecto de desambiguación, ha sido abordado a partir del relevamiento de técnicas de procesamiento de lenguaje natural, aprendizaje automático, y análisis de redes sociales, a partir de los datos extraídos de las producciones científicas. Los resultados parciales se presentan en las próximas secciones. Mientras que para el caso del método de recomendación se presentan los avances en el desarrollo de los componentes finales del mismo.

3 RESULTADOS Y OBJETIVOS

En este apartado se presentan los avances logrados en ambas líneas de trabajo.

3.1 MÉTODO DE DESAMBIGUACIÓN DE AUTORES

Se hizo un relevamiento de métodos de desambiguación actuales, considerando aspectos tales como: los datos que utilizan, las métricas para validar su funcionamiento y su impacto en los sistemas de información [9].

Uno de los métodos es utilizado por un repositorio académico actual, AMiner [1]. El mismo consiste en un *framework* basado en Campos Ocultos Aleatorios de Markov (o *Hidden Markov Random Fields*). Consiste en un modelo probabilístico que explora tanto las relaciones entre documentos como los atributos de cada documento.

Se formaliza el problema de desambiguación de la siguiente manera. Se tiene un conjunto de publicaciones, que representan las variables observables del Modelo de Markov. Dichos artículos están escritos por autores con un mismo nombre. Se aplica un método de *clustering* sobre el conjunto de *papers* en el que cada cluster corresponde a un autor diferente. Dado que se desconoce el número de autores, es por ello que el conjunto de autores representa el campo oculto aleatorio en el modelo de Markov. El mismo *framework* contiene un método de estimación del número de autores para el conjunto de publicaciones.

El objetivo a corto plazo en esta línea de trabajo es reproducir el algoritmo de AMiner para comprender mejor su funcionamiento y contrastarlo con un método de desambiguación a definir. Posteriormente se realizará un análisis de ambos métodos y en base a los resultados se determinará cuáles de estos métodos se adapta mejor para su implementación en el metabuscador.

3.2 MÉTODO DE RECOMENDACIÓN DE AUTORES

Tanto el esquema empleado para el almacenamiento de los datos de las entidades con las que opera el metabuscador, como una reseña parcial del proceso de recomendación de autores han sido presentados previamente [7], [8]. En lo que respecta al método de recomendación, en la publicación mencionada se detallaron cuestiones relativas a las fuentes de datos a utilizar, los aspectos que serían incluidos en la evaluación de los autores y las métricas asociadas a los mismos, dando lugar a la definición de las clases en función de las que serían catalogados y al algoritmo que efectúa la clasificación en sí misma.

A continuación, se presentan los componentes restantes del método, en particular aquellos que serán utilizados a partir del ingreso de la consulta del usuario en el metabuscador. Se propuso para el proceso de recomendación el uso de la técnica de filtrado basado en contenido [13], la cual plantea la generación

de recomendaciones a partir del estudio del contenido de los ítems a recomendar y las preferencias del usuario.

El proceso se compone de tres etapas sucesivas que se desarrollan a continuación:

[a] Para la etapa inicial, denominada de "**similaridad**" entre consulta y autores, se utilizaron los datos pertenecientes a las publicaciones de cada autor de la BD del SRI. Sobre los campos de título, resumen y palabras clave de cada publicación almacenada, se buscaron coincidencias con los términos de la consulta del usuario.

Resultando en un conjunto reducido de autores "*candidatos*" que se corresponden con los requerimientos del usuario. En este caso se simplificó el proceso seleccionando a todos los autores en cuyos perfiles se encontraron coincidencias con los términos de la consulta.

[b] Para la segunda etapa, que se denominó de "**relevancia**", el objetivo es determinar la relevancia de los autores candidatos para la recomendación. Para ello se establecieron los criterios, la técnica y las definiciones consideradas para la evaluación de los autores, los elementos utilizados y adaptados para el escenario propuesto, y la propuesta de un algoritmo para su aplicación.

Entre las técnicas aplicables para esta fase [14], [15], se determinó emplear un modelo de espacios vectoriales el cual permite desarrollar el siguiente postulado: "*los autores más relevantes para un término dado, pueden ser evaluados según la calidad de sus trabajos, para lo que existen métricas ampliamente aceptadas en la comunidad científica*". En función de ese principio, el proceso de filtrado toma en consideración dos factores: las métricas que determinan el impacto de un autor, siendo las mismas empleadas en el método de clasificación [8]; y métricas que estiman el peso de los términos de la consulta en los datos presentes en las publicaciones del autor. Para esto último se empleó el índice TF-IDF (*Term frequency – Inverse document frequency*) que es utilizado para este tipo de tareas en la literatura del área [16].

[c] Finalmente, en la fase de "**reparto**", se listan los ítems que resultan similares y relevantes para los usuarios. El desafío en esta etapa radica en la tarea de presentar el mayor número de ítems relevantes con el menor número de falsos positivos y por ello se busca la visibilidad de ítems con la mayor relevancia posible. Se propuso aplicar un método que permita emplear las clases de autores identificadas previamente [8].

Complementariamente, se utilizaron métodos de reparto como los empleados en elecciones gubernamentales, particularmente el de Sante-Laguë [17]. Considerando a las listas o partidos políticos equivalentes a cada una de las clases que pueden pertenecer los autores; y a la cantidad de votos de una lista o partido equivalente a la cantidad numérica de autores agrupados dentro de cada clase. Es decir, que la cantidad de autores de una clase definirá la cantidad de "*lugares*" que tendrá cada categoría en la presentación al usuario.

Estas operaciones se ejecutan sobre el conjunto de autores procesado en las etapas previas, resultando en un conjunto de autores a presentar al usuario como recomendación ante su consulta.

Actualmente el método completo se encuentra implementado y su fase de validación está siendo finalizada para ser presentado integralmente en una futura publicación.

4 FORMACIÓN DE RECURSOS HUMANOS

Este proyecto cuenta 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 Estímulo a las Vocaciones Científicas del Consejo InterUniversitario Nacional (CIN), tres tesis de maestría, dos de ellas enmarcadas en becas del Programa Estratégico de Formación de Recursos Humanos en Investigación y Desarrollo (PERHID) del CIN, y un trabajo de investigación posdoctoral con beca del Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Asimismo, la línea y el equipo de investigación se vinculan con investigadores de la Universidad de Castilla-La Mancha, España y de la Universidad de Sonora, México.

5 BIBLIOGRAFÍA

[1] J. Tang, "AMiner: Mining Deep Knowledge from Big Scholar Data," in *Proceedings of the 25th International Conference Companion on World Wide Web*, Geneva, Switzerland, 2016, pp. 373–373.

[2] 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.

[3] H. Li, I. Councill, W.-C. Lee, and C. L. Giles, "CiteSeerx: An Architecture and Web Service Design for an Academic Document Search Engine," in *Proceedings of the 15th International Conference on World Wide Web*, New York, NY, USA, 2006, pp. 883–884.

[4] H. Kuna, M. Rey, E. Martini, L. Solonezen, and L. Podkowa, "Desarrollo de un Sistema de Recuperación de Información para Publicaciones Científicas del Área de Ciencias de la Computación," *Rev. Latinoam. Ing. Softw.*, vol. 2, no. 2, pp. 107–114, 2013.

[5] H. D. Kuna, E. Martini, and M. Rey, "Evolución de un algoritmo de ranking para documentos científicos del área de las ciencias de la computación," XX Congreso Argentino de Ciencias de la Computación (Bs. As., 2014), 2014.

[6] H. Kuna, M. Rey, L. Podkowa, E. Martini, and L. Solonezen, "Expansión de Consultas Basada en Ontologías para un Sistema de Recuperación de Información," XVI Workshop de Investigadores en Ciencias de la Computación, 2014.

[7] 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.

[8] A. Cantero, M. Rey, and H. Kuna, "Clasificación de autores para un proceso de recomendación integrado a un metabuscador científico," in *XXIV Congreso Argentino de Ciencias de la Computación - Libro de Actas*, Tandil, Argentina, 2018.

[9] A. Canteros, E. Zamudio, and H. D. Kuna, "Desambiguación de autores para un sistema de recuperación de expertos en un contexto académico," in *XIX Simposio Argentino de Inteligencia Artificial (ASAI)-JAIIO 47 (CABA, 2018)*, CABA, Argentina, 2018.

[10] 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, 2015.

[11] Y. Qian, Q. Zheng, T. Sakai, J. Ye, and J. Liu, "Dynamic author name disambiguation for growing digital libraries," *Inf. Retr. J.*, vol. 18, no. 5, pp. 379–412, Oct. 2015.

[12] 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, 2017.

[13] R. Van Meteren and M. Van Someren, "Using content-based filtering for recommendation," in *Proceedings of the Machine Learning in the New Information Age: MLnet/ECML2000 Workshop*, 2000, pp. 47–56.

[14] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.

[15] P. Lops, M. de Gemmis, and G. Semeraro, "Content-based Recommender Systems: State of the Art and Trends," in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds. Boston, MA: Springer US, 2011, pp. 73–105.

[16] F. Cacheda, "Introduction to the Classic Models of Information Retrieval," *Rev. Gen. Inf. Doc.*, vol. 18, p. 365, 2008.

[17] V. R. González and A. L. Carmona, "Sistemas electorales basados en la representación proporcional," *eXtoikos*, no. 6, pp. 29–39, 2012.

Big Data como tecnología disruptiva en los nuevos modelos operativos organizacionales

Silvia Sánchez Zuaín, Lilia Palomo y Laura Sánchez Piccardi
Facultad de Ciencias Económicas y Facultad de Ciencias para la Innovación y el Desarrollo
Universidad Católica de Santiago del Estero
szuain@ucse.edu.ar - lilia.palomo@ucse.edu.ar - lsanchezpiccardi@gmail.com

RESUMEN

En la nueva era digital, el actual panorama empresarial se presenta con un alto grado de cambio y competitividad; los negocios y los modelos operativos globales deben ser ágiles y con capacidades adaptativas para responder a los retos de un entorno complejo y cambiante, pero que a la vez es generador de oportunidades de crecimiento. Estas oportunidades se apoyan, entre otros elementos fundamentales de competitividad, en las nuevas tecnologías como factor de innovación, junto al capital humano.

Las tecnologías disruptivas son aquellas que tienen como base la innovación (Big Data, virtualización, Cloud, ciberseguridad, realidad virtual, etc.) Tienen como denominador común también su capacidad de evolucionar rápidamente y adaptarse a diferentes sectores, generando nuevos modelos de negocio.

Un aspecto importante, de la reaceleración del cambio tecnológico, son los Big Data que están reconfigurando los negocios, transformando la organización interna y la arquitectura de las empresas. (Francisco Gonzales, 2014)

Actualmente, el éxito duradero depende de un modelo operativo de rápida adaptación, por lo tanto, las empresas tradicionales deben actuar como negocios que viven con la flexibilidad de velocidad y escala necesarias para seguir siendo relevantes sin importar qué interrupciones les espera.

Entonces, es necesario contar con un modelo operativo que pueda responder a las cambiantes condiciones del mercado y que permita interrumpir continuamente las estrategias de lanzamiento al mercado, propuestas de valor, y métodos con el fin de obtener competitividad.

De lo expuesto precedentemente, se desprende el propósito de este artículo sobre el Big Data como tecnología disruptiva dentro la economía digital, donde la integración de los negocios, las operaciones y la tecnología, son bloques de construcción de modelos operativos de nueva generación.

Palabras clave: Big Data, Tecnologías disruptivas, Modelo Operativo del Futuro, Modelo Operativo de Próxima Generación; Economía digital.

CONTEXTO

El presente artículo se presenta como una investigación exploratoria, del Big Data como tecnología disruptiva, ante la necesidad de las organizaciones modernas de contar con nuevos modelos operativos para apoyar las cambiantes condiciones del mercado y garantizar así su supervivencia en la nueva era digital.

Integra una línea del trabajo de investigación de cátedra, "Hacia una economía digital: integración de los negocios, las operaciones y la tecnología en nuevos modelos operativos para el futuro", perteneciente a la Facultad de Ciencias Económicas, de la Universidad Católica de Santiago del Estero.

Este tipo de actividad, promueve la interacción vertical y horizontal, a partir de las asignaturas: Sistemas y Tecnología de la Información (carrera Contador Pùblico) de la Facultad de Ciencias Económicas; Ingeniería de Software (carrera de Ingeniería en Informática) y Análisis Numérico (carrera de Electrónica) de la Facultad de Ciencias para la Innovación y el Desarrollo.

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

Históricamente, el desarrollo económico fue impulsado por la dinamización de los factores de producción, mientras que en los últimos años, se sumó la intensificación en el uso de las tecnologías. El impacto tecnológico que desarrolla la economía digital son las nuevas capacidades de la infraestructura de Internet, y las potencialidades de las tecnologías emergentes como: Cloud, Big Data, Internet de todas las Cosas (IoE), Smart Cities, o Industria 4.0.

Particularmente, la tecnología digital tiene un impacto decisivo en las organizaciones de todos los tamaños y sectores, en formas poderosas y profundas, cuestionando paradigmas clave de gestión y cambiando, en definitiva, el concepto actual de negocio.

Se está viviendo una reaceleración del cambio tecnológico. Uno de los aspectos importantes de este cambio es cómo los Big Data están reconfigurando los negocios, transformando la organización interna y la arquitectura de las empresas. Philip Evans (Boston Consulting Group) describe los dos motores de las tecnologías de la información (TI) que están remodelando la organización, las estrategias y las estructuras empresariales: una es la deconstrucción de las cadenas de valor y la otra la polarización de las economías de masa. (Philip Evans, 2015)

Los modelos de negocio evolucionan, en parte debido al profundo impacto de las tecnologías de la información (TI). En una proyección a largo plazo, el profesor Haim Mendelson (Stanford Business School) prevé que las TI proseguirán su mejora continua de rendimiento. El efecto combinado de las tecnologías móviles, los dispositivos y sensores, la computación en la nube y las tecnologías de los Big Data refinarán la estructura de los modelos de negocio futuros. Mendelson considera que las formas tradicionales de innovación marcarán la diferencia. Primero, porque los modelos de negocio subyacentes requerirán innovación continua y, segundo, porque los nuevos productos seguirán precisando modos de innovación convencionales. (Haim Mendelson, 2015)

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación del presente trabajo tiene como eje central el Big Data como una tecnología disruptiva, que causa gran impacto en la economía empresarial, ante la necesidad de contar con nuevos modelos operativos para apoyar las cambiantes condiciones del mercado y garantizar así su supervivencia en la nueva era digital.

Se pueden mencionar los siguientes supuestos que dan estructura a la temática o campo de estudio del proyecto:

- Se utiliza el término disruptivo para hablar de aquellas tecnologías que causan un gran impacto en el ámbito empresarial.
- Big Data, es un concepto disruptor, que está reconfigurando los negocios, transformando la organización interna y la arquitectura de las empresas.
- La empresa necesita contar con un Modelo Operativo Digital, que es “una nueva forma de administrar la organización que combina tecnologías digitales y capacidades de operaciones de manera integrada y bien secuenciada para lograr mejoras en los cambios en los ingresos, la experiencia del cliente y el costo" (McKinsey Digital, 2017).
- Para competir con los nativos digitales, las empresas deben reorganizarse en torno a los trayectos del cliente y combinar capacidades con tecnologías digitales.
- El crecimiento exponencial de la tecnología, se estima que para el 2020 habrá 7 veces más dispositivos en red que personas en el mundo (Luis Enrique Arce, 2017)
- Los Big Data son, sin duda, uno de los ingredientes claves para la transformación exitosa del modelo

operativo organizacional.

- Ningún ámbito de la actividad humana ni sector de la industria será inmune a la total reorganización que traen los Big Data. (Kenneth Cukier, 2015)
- Este nuevo mundo de datos afectará a dos áreas de las políticas públicas: empleo (las tecnologías generarán a medio plazo una oleada de desempleo estructural) y privacidad (porque los datos se recogen de manera invisible y pasiva, como subproducto de otro servicio).

Se trabajará en situar el Big Data como tecnología disruptiva, donde la integración de los negocios, las operaciones y la tecnología son bloques de construcción de modelos operativos de próxima generación.

3. RESULTADOS ESPERADOS

El objetivo de esta línea de investigación plantea realizar un estudio y análisis del Big Data, como un elemento disruptivo, en la integración de los negocios, las operaciones y la tecnología de los nuevos modelos operativos del futuro, que permitan a las organizaciones, de todos los tamaños y sectores, apoyar las cambiantes condiciones del mercado y garantizar así su supervivencia en la nueva era digital.

Para el logro del objetivo general, este trabajo se centrará en unos objetivos más específicos**:**

- Relevar y examinar los diferentes enfoques técnicos y estratégicos disponibles actualmente con el fin de aumentar agilidad a las empresas.
- Identificar las limitaciones existentes que impidan a las empresas ser realmente disruptivas en nuevas áreas a la vez que apoyan sus líneas tradicionales de negocios.
- Identificar las características esenciales que las empresas deben adoptar e integrar en sus modelos operativos para transformarse para el futuro y garantizar su existencia continua en un mundo disruptivo.

Las actividades que se llevarán a cabo son las siguientes:

- Situar el Big Data dentro de las cinco dimensiones del Modelo Operativo del Futuro: procesos, estructura y gobernanza, ecosistemas, tecnología, capacidades y cultura. Donde, una de esas capacidades, es el procesamiento autónomo de datos que utiliza herramientas sofisticadas para descubrir ideas y hacer recomendaciones.
- Identificar el Big Data, como facilitador, dentro de las categorías de habilitadores digitales que permiten afrontar la digitalización de las organizaciones (MITYC, 2015).
- Presentar una foto global, desde el punto de vista técnico, de las tecnologías disruptivas (tecnologías TIC), que aplicadas en toda la cadena de valor, están impulsando beneficios de alto impacto en las empresas, a través de una gama de aplicaciones.

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. Cabe destacar, la incorporación de un alumno en carácter de becario, para motivarlo a realizar su trabajo final de grado en el área de este proyecto

El grupo hace difusión y formación de recursos humanos desde las asignaturas: Sistemas y Tecnología de la Información, Análisis Numérico, e Ingeniería de Software.

5. BIBLIOGRAFIA

- Francisco Gonzales, 2014. "Reinventar la empresa en la era digital". *BBVA*. Disponible en: https://www.bbvaopenmind.com/wp-content/uploads/2015/01/BBVA-OpenMind-libro-Reinventar-la-Empresa-en-la-Era-Digital-empresa-innovacion1-1.pdf
- Haim Mendelson, 2015. "Modelos de negocio, tecnologías de la información y la empresa del futuro". *BBVA*. Disponible en: https://www.bbvaopenmind.com/articulos/modelos-de-negocio-tecnologias-de-la-informacion-y-la-empresa-del-futuro/
- Philip Evans, 2015. "De la deconstrucción a los big data: cómo la tecnología está transformando las empresas". *BBVA*. Disponible en: https://www.bbvaopenmind.com/articulos/de-la-deconstrucion-a-los-big-data-como-la-tecnologia-esta-transformando-las-empresas/
- McKinsey Digital (2017). Introducing the next-generation operating model. Disponible en: www.mckinsey.com/business-functions/digital-mckinsey/our-insights/introducing-the-next-generation-operating-model
- Luis E. Arce, 2017. "Nuevos modelos de Negocio. Nuevas Capacidades Organizacionales. Nuevas Formas de jugar en una economía digital". Disponible en: https://computerworld.com.ec/files/CWBIDENRIQUEARCEQUITO-ilovepdf-compressed.pdf
- Kenneth Cukier, 2015. "Los big data y el futuro de los negocios". *BBVA*. Disponible en: www.bbvaopenmind.com/articulos/los-big-data-y-el-futuro-de-los-negocios/
- MITYC (2015). Ministerio de Industria, Energía y Turismo. Informe preliminar: "Industria Conectada 4.0: la transformación digital de la industria española". España. Disponible en: http://www6.mityc.es/IndustriaConectada40/informe-industria-conectada40.pdf
- Deloitte, 2017. Digital era Technology Operating Models. Disponible en: www2.deloitte.com/content/dam/Deloitte/nl/Documents/technology/deloitte-nl-digital-era-tom-v1.pdf
- General Electric (2012). "Industrial Internet: Pushing the Boundaries of Minds and Machines". Disponible en: www.ge.com/docs/chapters/Industrial_Internet.pdf
- Karel Dörner y otros (2015). What Digital Really Means. McKinsey Digital. Disponible en: www.mckinsey.com/industries/high-tech/our-insights/what-digital-really-means
- Marc Brown (2019). "Six Digital Transformation Barriers CMOs Need to Address". Disponible en: which-50.com/six-digital-transformation-barriers-cmos-need-to-address/
- McKinsey Digital (2018). Digital reinvention: Unlocking the 'how'. Disponible en: www.mckinsey.com/~/media/McKinsey/Business%20Functions/McKinsey%20Digital/Our%20Insights/Digital%20Reinvention%20Unlocking%20the%20how/Digital-Reinvention_Unlocking-the-how.ashx
- Siemens (2016). "España 4.0. El reto de la transformación digital de la economía". Disponible en: w5.siemens.com/spain/web/es/estudiodigitalizacion/Documents/Estudio_Digitalizacion_Espana40_Siemens.pdf
- Tanguy Catlin (2018).Digital strategy: The four fights you have to win. Digital McKinsey&Company. Disponible en: www.mckinsey.com/business-functions/digital-mckinsey/our-insights/digital-strategy-the-four-fights-you-have-to-win
- World Economic Forum (2016). Digital Transformation of Industries: Digital Enterprise. Disponible en: http://reports.weforum.org/digital-transformation/wpcontent/blogs.dir/94/mp/files/pages/files/digital-enterprise-narrative-final-january-2016.pdf

Ciencia de datos en bibliotecas populares sanjuaninas

Prog. Luis Olguin, Mag. Alejandra Malberti, Mag. Raúl Klenzi
Instituto de Informática – Departamento de Informática
Facultad de Ciencias Exactas, Físicas y Naturales Universidad Nacional de San Juan
lolguin@iinfo.unsj.edu.ar; amalberti@gmail.com; rauloscarklenzi@gmail.com

RESUMEN

Las *Bibliotecas Populares* son creadas, gestionadas y mantenidas por los propios asociados. Son entidades autónomas (ONGs) que adaptan sus acciones a las necesidades de su público. Existen en Argentina desde hace más de 150 años, y han ido adecuándose a la situación social, política y económica del país. Son, sin lugar a dudas, un lugar de "acceso abierto" al conocimiento, asumiendo el papel de mediador entre el saber y el ciudadano común.

Actualmente las bibliotecas populares de San Juan, que poseen sistemas de gestión automatizada, no elaboran estrategias de acción que fortalezcan su accionar **analizando los datos** que el sistema de gestión genera, perdiendo una oportunidad de descubrir particularidades de sus visitantes que les permitan mejorar su posición en el contexto social.

La presente propuesta, encarada a través del proyecto *Ciencia de Datos en Bibliotecas Populares Sanjuaninas (CiDaBi-SJ),* pretende abordar *el paradigma de la Ciencia de Datos* centrándose en el análisis, extracción de conocimiento y valor asociado al conjunto de datos generados por los distintos sistemas de gestión bibliotecaria que poseen las bibliotecas populares de San Juan.

Se indaga en cómo adaptar la *analítica de datos* a las necesidades de una biblioteca popular, estudiando cada caso y aportando soluciones realistas.

***Palabras clave**:* Ciencia de Datos, Biblioteca Popular, Visualización de Datos

CONTEXTO

Esta investigación se enmarca dentro de *la convocatoria de Vinculación Tecnológica de la Secretaría de Políticas Universitarias "Agregando Valor 2017" (RS-2017-31806560-APN-SECPU#ME)* y forma parte de las líneas abordadas por *e*l *Laboratorio de Sistemas Inteligentes para la Búsqueda de Conocimiento en Datos Masivos* y el *Laboratorio de Bibliotecas Digitale*s del Instituto de Informática de la FCEFN (*Resol. 2/2015-CD-FCEFN).* Los docentes que integran el equipo de CiDaBi-SJ vienen trabajando en temáticas relacionadas desde el año 2005. Cómo antecedentes se encuentran los proyectos *"Descubrimiento de conocimiento a través de Data Warehousing y Data Mining, en los datos de la Biblioteca de la Facultad de Ciencias Exactas, Físicas y Naturales",* 21/E639, desarrollado en el período 2005-2007*; "Búsqueda estratégica de conocimiento en los datos de biblioteca y de alumnos de la FCEFN",* 21/E824 , período 2008-2010; *"Minería de datos en la determinación de patrones de uso y perfiles de usuarios",* 21/E889, periodo 2011-2013*, "Extracción de Conocimiento en Datos*

Masivos, periodo 2014-2015 y "Visualización y Deep Learning en Ciencia de Datos", periodo 2018-2019.

1. INTRODUCCIÓN

El trabajo que llevan adelante las bibliotecas populares sanjuaninas, que poseen sistemas de gestión automatizados incluye la aplicación de *técnicas de análisis de la información* que estos sistemas generan - en general por falta de formación de RRHH en estas temáticas- encontrándonos con *bibliotecarios operadores del sistema*, es decir adaptados a lo que los diferentes módulos ofrecen (catalogación, circulación, asociación, etc.) sin tener como objetivo la *búsqueda de patrones ocultos y útiles o resultados estadísticos novedosos*.

La acción de usar un sistema de gestión bibliotecaria, no implica la extracción de patrones automáticamente y por tanto se desperdicia la oportunidad de usar esas *pepitas de información* para la mejora de otros servicios como ser la *diseminación selectiva de la información* entre los asociados a una biblioteca popular (Ruiz-Lobaina, Esther Marina y Pedro Lázaro Romero-Sánchez. 2017).

En la actualidad *parece natural* que las bibliotecas exploren el uso de las herramientas de ciencia de datos para satisfacer la demanda creciente de sus usuarios, tal es el caso de la biblioteca de Signapure que haciendo uso de herramientas de *big data* ha logrado optimizar los procesos de selección y oferta de obras a sus usuarios, tarea que antes se efectuaba solo a *criterio humano*. Esto se logra a través del *análisis de demanda*, técnica que permite pronosticar la demanda de los usuarios, aumentando en consecuencia la satisfacción del asociado (SINGAPORE LIBRARY MINES BIG DATA, 2014).

El análisis de grandes datos sin dudas es un paso que los profesionales de las bibliotecas deben dar para poder extraer información valiosa de sus repositorios (Alonso Arévalo, Julio; Vázquez Vázquez, Marta. 2016).

En San juan lamentablemente no se aborda, desde el Instituto de Formación en Bibliotecología dependiente de Ministerio de Educación de la provincia, la formación de RRHH en estas técnicas, por lo que el abordaje de la temática a través de la experiencia del equipo de trabajo de IdeI sin duda será un aporte para las bibliotecas populares involucradas.

En la actualidad el enunciado "los datos masivos están en el punto de remodelar nuestro modo de vivir, trabajar y pensar" (Mayer-Schönberger, 2013) debe ser analizado seriamente por las bibliotecas y, dado que el conocimiento (o parte de él) se almacena en ellas, contar con herramientas que permitan extraer *información de calidad, oportuna y adecuada* es una necesidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La posibilidad de almacenar enormes volúmenes de datos exige el estudio de nuevas formas de representarlos para extraer información valiosa. En el proyecto CiDaBi-SJ se propone el estudio de herramientas y técnicas que permitan analizar grandes volúmenes de datos y generar representaciones visuales que posibiliten la toma de decisiones tanto a los bibliotecarios

como a los usuarios finales de una biblioteca popular.

La línea principal de investigación, que se ajusta a una investigación en informática aplicada, consiste en el empleo de soluciones, especialmente software libre, para la toma de decisiones en el ámbito de una biblioteca popular.

3. RESULTADOS OBTENIDOS / ESPERADOS

El proyecto Ciencia de Datos en Bibliotecas Populares Sanjuaninas (CiDaBi-SJ) es un proyecto de investigación aplicada que se sustenta en la hipótesis que "las bibliotecas populares deben asumir la aplicación de la Ciencia de Datos para obtener una mejora en la calidad de sus servicios y procesos que les permita mejorar su competividad en un marco de sostenibilidad y actuación socialmente responsable".

Para llevar a cabo la hipótesis se precisará definir indicadores y variables que ayuden a establecer si es factible la aplicación de las técnicas de ciencia de datos en una biblioteca popular.

Se espera que al finalizar este proyecto sea posible mostrar patrones de uso de la biblioteca popular por parte de los usuarios que a ella acuden desde el enfoque de la ciencia de datos, y que estos resultados sean interpretados de manera correcta tanto por parte de los bibliotecarios como del público en general, es decir los destinatarios de esta investigación.

La ejecución del proyecto CiDaBi-SJ permitirá a los alumnos participantes adiestrarse en los procesos de extracción de datos independientemente de su fuente, la limpieza de los datos eliminando los atributos que no aportan valor a los resultados buscados, dominar técnicas de procesamiento de datos usando diferente métodos y algoritmos de minería de datos, y la visualización gráfica (cualitativa y cuantitativamente) de los datos y resultados.

A los fines de alcanzar los objetivos propuestos para CiDaBi-SJ se prevé como eje principal la participación activa de la entidad adoptante, tomando en cuenta los medios y recursos disponibles. Se buscará gestionar de manera sistemática el compromiso de la entidad adoptante para encontrar un balance entre sus expectativas y los objetivos del proyecto. Los estudiantes se verán favorecidos por el nuevo conocimiento además de la posibilidad concreta de transferencia al medio. Las temáticas abordadas en el proyecto brindarán a los docentes participantes nuevas líneas de investigación a abordar, así como la posibilidad de realizar aportes al conocimiento en las áreas relacionadas.

4. FORMACIÓN DE RRHH

Olguín, Luis Alberto. Integrante del proyecto. Es actualmente maestrando de la Maestría en Informática de la Universidad Nacional de San Juan, siendo el título de su trabajo de tesis Red de Co-Préstamo en Bibliotecas

Formación de alumnos avanzados de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, en el uso de herramientas de la ciencia de datos.

*Formación de docente*s de las carreras del Departamento de Informática, integrantes del

proyecto, en el tratamiento de datos relativos a la problemática abordada.

Entrenamiento a los RRHH de la entidad adoptante en el uso de herramientas informáticas, por medio de las capacitaciones previstas en el plan de trabajo de CiDaBi-SJ.

5. **REFERENCIAS**

Alonso Arévalo, Julio; Vázquez Vázquez, Marta (2016). *Big Data: la próxima «gran cosa» en la gestión de la información* - . BiD: textos universitaris de biblioteconomia i documentació , núm. 36 (jun) . <http://bid.ub.edu/es/36/alonso.htm>. [Consulta: 10-02-2019].

Mayer-Schönberger, (2013). *Big Data: La revolución de los datos masivos.* Madrid: Turner. <http://catedradatos.com.ar/media/3.-Big-data.-La-revolucion-de-los-datos-masivos-Noema-Spanish-Edition-Viktor-Mayer-Schonberger-Kenneth-Cukier.pdf>. [Consulta: 10-02-2019].

Ozdemir, S. (2016). *Principles of Data Science*. Packt Publishing Ltd.

Ruiz-Lobaina, Esther Marina y Pedro Lázaro Romero-Sánchez, (2017). *Búsqueda de patrones para mejorar productos y servicios en las bibliotecas*. Investigación Bibliotecológica: Archivonomía, Bibliotecología e Información 72 (31): 209-225. <http://dx.doi.org/10.22201/iibi.0187358xp.2017.72.57830>.[Consulta: 10-02-2019].

Singapore Library Mines Big Data (2014). Business Information, ASEAN, vol. 2. <http://www.asiandatascience.com/wp-content/uploads/2014/09/BI_Asean_Sept_2014_FINAL.pdf> [Consulta: 08-02-2019].

Descubrimiento de Conocimiento en Bases de Datos

Lautaro Ramos, Esteban Schab, Ramiro Rivera, Cristhian Richard, Patricia Cristaldo, Juan Pablo Núñez, Giovanni Daián Rottoli, Juan Manuel Ríos, 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
{ramosl, schabe, riverar, richardc, cristaldop, nunezjp, rottolig,
riosj, retamars, casanovac,debattistaa}@frcu.utn.edu.ar

Leticia Cagnina, Norma Edith Herrera
Departamento de Informática, Universidad Nacional de San Luis, San Luis, Argentina
{lcagnina,nherrera}@unsl.edu.ar

Resumen

En la actualidad se generan diariamente grandes cantidades de datos de diversos tipos (e.g. textos, imágenes, audios y videos) generando nuevas fuentes de información que pueden ser aprovechadas para agregar valor al trabajo de las organizaciones. Particularmente el análisis automático de textos (análisis de sentimientos, minería de opinión) ha ganado terreno como alternativa o complemento a las fuentes de datos tradicionales de información de las organizaciones, cobrando relevancia las técnicas de Minería de Textos. La mayoría de los algoritmos, herramientas y recursos disponibles para Minería de Textos han sido probados y/o desarrollados para el idioma inglés, y por tanto presentan dificultades al ser empleados sobre textos escritos en otros idiomas como el español. Es por esta razón que es necesario trabajar en la elaboración de recursos específicos y en la adaptación de algoritmos y herramientas que contemplen las particularidades del idioma español con el fin de poder conseguir resultados de mayor calidad.

En este artículo se presentan los tópicos de interés del proyecto *Descubrimiento de Conocimiento en Bases de Datos*, en el que se investigan técnicas de minería de textos aplicables al procesamiento de textos en lenguaje español. En particular, se realizará el estudio, análisis y comparación de algoritmos de minería de textos utilizando corpus de textos en lenguaje español, para posteriormente proponer adaptaciones o mejoras a los mismos. Asimismo, se pretende evaluar el desempeño de técnicas de minería de datos sobre conjuntos de datos tradicionales complementados con información extraída a partir de textos relacionados.

Palabras clave: minería de datos, minería de textos, bases de datos, descubrimiento de conocimiento, idioma español.

Contexto

El presente trabajo se desarrolla en el ámbito del proyecto *Descubrimiento de conocimiento en Bases de Datos (Código 5109)* 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.

Para aprovechar el conocimiento potencial que puede obtenerse a partir del análisis automático de textos surge la Minería de Textos, consistente en el proceso de extraer patrones relevantes a partir de un gran conjunto de textos con el propósito de obtener conocimiento. Abarca una serie de tareas que se enfocan en distintos aspectos del análisis de textos. Algunas de las más relevantes [1]:

- Recuperación de información (*Information Retrieval*, IR): es la tarea de encontrar material de naturaleza no estructurada (generalmente textos) proveniente de grandes colecciones que satisfagan determinadas necesidades de información [2]. Una tarea crucial para un sistema de IR es indexar la colección de documentos para hacer que sus contenidos sean accesibles de manera eficiente. Generalmente la indexación se realiza sobre una representación lógica del documento, que puede consistir en un conjunto de palabras clave o términos relevantes que aparezcan en el texto [3].
- Procesamiento del Lenguaje Natural: es un campo de las ciencias de la computación que combina Inteligencia Artificial y conceptos lingüísticos con el fin de hacer que oraciones o palabras escritas en lenguaje natural puedan ser interpretados por programas de computadoras [1, 4].
- Resumen de textos (*Text Summarization*): es la tarea de producir un resumen conciso y fluido preservando el contenido clave de la información y el significado general de una colección de textos [6].
- Extracción de Información (*Information Extraction*, IE): es una subdisciplina de la Inteligencia Artificial que se aboca a la identificación, y consecuente clasificación y estructuración en grupos semánticos, de información específica que se encuentran en fuentes de datos no estructurados, como el texto en lenguaje natural, lo que hace que la información sea más adecuada para las tareas de procesamiento de la información [5].
- Métodos de Aprendizaje Supervisado y No Supervisado: los métodos de aprendizaje supervisado son técnicas de aprendizaje automático relacionadas con entrenar un modelo, por ejemplo, de clasificación, utilizando un conjunto de datos de entrenamiento para realizar predicciones sobre datos desconocidos de antemano. Existe una amplia gama de métodos supervisados, como clasificadores de vecinos más cercanos, árboles de decisión, clasificadores basados en reglas y clasificadores probabilísticos.

Los trabajos sobre minería de textos en español que se presentan en la actualidad se enfocan principalmente en Análisis de Sentimientos o Minería de Opinión, en los cuales se evidencian dos enfoques: uno basado en el empleo de lexemas, y otro en técnicas de *Machine Learning*, para identificar los sentimientos expresados en los textos. En la gran mayoría de estos trabajos se utilizan recursos traducidos de forma automática generados para otros idiomas, como el inglés, lo cual manifiesta una escasez de recursos genuinos para el lenguaje español.

Existen también trabajos sobre perfilado de autor, en los que se menciona la dificultad de encontrar colecciones de textos adecuadamente etiquetados y con poco ruido [7]. A partir de eso se han producido trabajos tendientes al desarrollo de conjuntos de textos en español específicos para esta tarea [8,9].

Las técnicas de minería de datos de texto se han propuesto para tareas de estudios bibliográficos, simplificación de textos y etiquetado semántico. Se ha propuesto la aplicación de algoritmos de clasificación [10] para identificar automáticamente el dominio disciplinar de un nuevo texto académico en un repositorio bibliográfico mediante la construcción de lexemas compartidos en cada disciplina.

2. Líneas de Investigación, Desarrollo e Innovación

La línea de trabajo principal de nuestro proyecto de investigación es el estudio,

análisis y comparación de técnicas de minería de datos para el tratamiento de textos en español, el análisis de su desempeño en distintos escenarios y la propuesta de modificaciones o mejoras a las técnicas de minería de textos existentes para incrementar la calidad de los resultados en el tratamiento de textos en español. También está previsto realizar una evaluación del funcionamiento de técnicas de minería de datos sobre conjuntos "tradicionales" enriquecidos con atributos provenientes de textos relacionados.
Se espera poder realizar vinculaciones con empresas u organizaciones que puedan obtener beneficios de la aplicación de técnicas de minería de textos en español.

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 [11], 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 [12], utilizando scripts desarrollados en lenguaje R [13]. 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. En la actualidad, se está trabajando en el denominado Modo 2, estudiando instituciones de pertenencia de los artículos provenientes de Scopus y además de 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 [14].
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 [15].

2.3. 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 metalmecánico 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.

2.4. Visualización de datos

La generación y el almacenamiento de grandes volúmenes de información hacen que el mismo pase desapercibido y muchas veces se pierda 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 [16, 17].

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[18], Gephi[19], D3js[20], React D3[21] y Shiny[22].

2.5. 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. [23, 24] Partiendo de la propuesta metodológica de CRISP-DM [25] se realizó una adaptación empleando dichos fundamentos ágiles. Se formalizó dicha adaptación como una propuesta de metodología ágil para la gestión de proyectos de ciencia de datos. [26, 27]

3. Resultados obtenidos y esperados

Con este proyecto se espera lograr aplicaciones novedosas de técnicas y herramientas de minería de textos para textos en español, en particular en áreas de estudio como bibliometría y la teoría de establecimiento de agenda. Estas iniciativas se desarrollan mediante la aplicación de la metodología ágil para proyectos de ciencias de datos propuesta. [26, 27]

4. Formación de Recursos Humanos

En el marco del proyecto se están desarrollando dos tesis de maestría y dos de sus integrantes están cursando carreras de doctorado. Se cuenta con un becario graduado con beca de iniciación a la investigación y dos becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación. Está prevista la realización de al menos dos Prácticas Supervisadas en el marco del proyecto.

5. Referencias

[3] Allahyari, Mehdi et al. A Brief Survey of Text Mining: Classification, Clustering and Extraction Techniques. 2017. arXiv:1707.02919v2

[4] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to information retrieval. Cambridge University Press, 2008.

[5] S. Ceri, A. Bozzon, M. Brambilla, E. Della Valle, P. Fraternali, and S. Quarteroni, The Information Retrieval Process. In Web Information Retrieval, Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 13–26.

[6] D. Khurana, A. Koli, K. Khatter, and S. Singh, “Natural Language Processing: State of The Art, Current Trends and Challenges,” Aug. 2017.

[7] Moens, Marie-Francine. Information Extraction: Algorithms and Prospects in a Retrieval Context. Springer Netherlands, 2006.

[8] M. Allahyari et al. Text Summarization Techniques: A Brief Survey. Jul. 2017.

[9] M. J. Garciarena Ucelay, M. P. Villegas, L. Cagnina, and M. L. Errecalde. Cross domain author profiling task in spanish language: an experimental study. *JCS&T*, vol.

15, no. 02, pp. p. 122-128. Nov. 2015.

[10] M. P. Villegas, M. J. Garciarena Ucelay, M. L. Errecalde, and L. Cagnina. A Spanish text corpus for the author profiling task. 2014.

[11] M. J. Garciarena Ucelay and M. P. Villegas. Determinación del Perfil del Autor de Documentos en Español. Universidad Nacional de San Luis, 2015.

[12] R. Venegas. Clasificación de textos académicos en función de su contenido léxico-semántico. Rev. signos, vol. 40, no. 63, pp. 239–271, 2007.

[13] SCOPUS. http://www.scopus.com Accedido 03/2018.

[14] Scopus API. https://goo.gl/mqpFpA Accedido 03/2017.

[15] R Project. https://www.r-project.org/ Accedido 03/2018.

[16] M. McCombs and D. Shaw. The agenda-setting function of mass media. Public opinion quarterly, 36(2):176–187, 1972.

[17] Yeooul 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.

[18] Sadiku, Matthew (2016). Data Visualization. International Journal of Engineering Research And Advanced Technology(IJERAT).Volume. 02. Issue.12. p. 11-16.

[19] 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).

[20] Tableau. https://www.tableau.com/es-es Accedido 03/2018.

[21] Gephi. https://gephi.org/ Accedido 03/2018.

[22] Data-Driven Documents. https://d3js.org/ Accedido 03/2018.

[23] React D3. http://www.reactd3.org/ Accedido 03/2018.

[24] Shiny from R Studio. https://shiny.rstudio.com/ Accedido 03/2018.

[25] Ken Schwaber and Jeff Sutherland. The scrum guide. Scrum Alliance, 2011, vol. 21.

[26] Manifesto for Agile Software Development. Agile Alliance. https://goo.gl/xRFCVL . Accedido 03/2018.

[27] Chapman, Clinton, Kerber, Khabaza, Reinartz, Shearer, & Wirth. CRISP-DM 1.0 Step-by-step data mining guide. 2000.

[28] Cristaldo, Richard, Rivera, Schab, Anabella De Battista. Propuesta Metodológica de Enfoque "Híbrido" para la Gestión de Proyectos de Minería de Datos. SABTIC 2018. ISSN 2237-2970

[29] Cristaldo, Schab, Richard, Rivera, Anabella De Battista, Retamar, Herrera. Adecuación de una Propuesta Metodológica de Enfoque "Híbrido" para la Gestión de Proyectos de Ciencia de Datos. 6to CoNaIISI. 29 y 30 de Noviembre de 2018 – Universidad CAECE – Mar del Plata, Bs. As., Argentina.

El big data desde la perspectiva de sus implicaciones jurídicas en la evidencia digital

Lilia Palomo, Laura Sánchez Piccardi y Sergio Mario Guillet
Facultad de Ciencias Políticas, Sociales y Jurídicas - Universidad Católica de Santiago del Estero
lilia.palomo@ucse.edu.ar – lsanchezpiccardi@gmail.com – smguillet@yahoo.com.ar

RESUMEN

La sociedad moderna se encuentra cada vez más tecnológica, lo que conlleva la aparición de un ecosistema digital muy diverso. En él conviven aplicaciones informáticas en entornos corporativos, bases de datos de información pública y privada, redes sociales y multitud de terminales desde los que se puede acceder, producir nuevos datos y extraer información. Los datos generados por estas aplicaciones están en todas partes, son de múltiple naturaleza y se almacenan en distintos lugares y formatos. El conjunto de toda esta explosión de información recibe el nombre de Big Data ("Datos Masivos" o "Macrodatos") y, por extensión, así también se denomina al conjunto de herramientas, técnicas y sistemas destinados a extraer todo su valor [1].

Como pasó con la revolución científica y como ocurre con cualquier innovación, la gestión de datos basada en el nuevo paradigma del Big Data carece, por sí mismo, de un tamiz ético que permita juzgar los avances que pueda conllevar, o los medios necesarios para conseguirlos.

Por lo tanto, también parece natural que los conflictos requieran cada vez en mayor medida de análisis y certificaciones digitales de los datos que se aportan a un procedimiento de resolución de disputas, sea judicial o arbitral.

Una investigación digital exitosa deberá identificar todas las fuentes de datos relevantes y abordar el acceso y análisis de todos esos datos. En última instancia deberá ofrecer conclusiones objetivas y sustentadas en una argumentación lógica y repetible. Lo que pone de manifiesto la necesidad de conocimiento y manejo profesional en lo atinente a lo probatorio.

Por esta razón, este proyecto propone una investigación exploratoria cualitativa donde se analizan las condiciones y estrategias del uso del Big Data en el contexto de la evidencia digital y su incidencia en el proceso de investigación judicial, siguiendo de cerca las políticas y legislación vigentes, a fin de obtener una Guía de Referencia con el fin de mejorar el tratamiento del Big Data, en el marco de la evidencia digital.

Palabras clave: Big Data, evidencia digital, prueba digital, prueba electrónica, forensia, investigación judicial.

CONTEXTO

La facultad de Ciencias Políticas, Sociales y Jurídicas perteneciente a 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, en el año 2018, desde las asignaturas Informática de la facultad de Ciencias Políticas, Sociales y Jurídicas, Ingeniería de Software y Análisis Numérico de la Facultad de Ciencias para la Innovación y el Desarrollo, surge el proyecto "Implicancias jurídicas del big data en la evidencia digital y su incidencia en el proceso judicial", 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.

El presente artículo presenta una línea de investigación que surge de ese proyecto, ante la necesidad en esta nueva era digital, de analizar las condiciones y estrategias del uso del Big Data desde la perspectiva de sus implicaciones jurídicas en la evidencia digital, siguiendo de cerca las políticas y legislación vigentes.

1. INTRODUCCIÓN

De manera ineludible, las tecnologías de la información y la comunicación (TICs) han irrumpido en todos los ámbitos y el proceso judicial no es la excepción. En el contexto donde un proceso está cada vez más embebido por las nuevas tecnologías, la gestión probatoria informática se torna una pieza clave para el correcto funcionamiento de las estrategias profesionales. Y para ello es imprescindible el manejo adecuado de esta área temática, en donde se interceptan lo jurídico y lo técnico [2].

Actualmente, en la era de la desmaterialización del documento, incrementándose la actuación a través de medios electrónicos y la progresiva desaparición del papel; las personas alrededor de todo el mundo descargan, crean, intercambian información e interactúan en la red [3]. Este impacto de la tecnología y la evolución hacia una sociedad cada vez más digital pone de manifiesto que se ha vuelto imprescindible la aportación en los juicios de las llamadas *evidencias digitales*[1].

Los datos tienen un alto valor, dado que la base de las demandas y de la acreditación de prueba se basa, cada vez en mayor medida, en datos en formato electrónico; se está volviendo cada vez más relevante la capacidad de analizar y explicar las acciones que llevan a cabo individuos, empresas e instituciones por medio de sus sistemas de información automatizados.

[1] Se entiende por evidencia digital a los datos que constan en formato electrónico y que constituyen elementos de prueba, comprendiendo las etapas de extracción, procesamiento e interpretación [4].

También como pasó con la revolución científica y el nuevo conocimiento, la gestión de datos basada en Big Data carece, por sí mismo, de un tamiz ético que permita juzgar los avances que pueda conllevar, o los medios necesarios para conseguirlos [5]. Y lo que anteriormente era una herramienta de modernización en las comunicaciones y en el sistema organizacional se convierte en una fuente del delito.

Además, las pruebas digitales son cada día más utilizadas en los juzgados sin que, al menos de momento, haya una doctrina que clarifique en qué supuestos es válida, dado la gran facilidad de manipulación que entrañan debido en parte a los peligros de suplantación o a la modificación del terminal originario del mensaje hace que su uso no resulte fácil [6].

Es tarea del derecho procesal encontrar el medio para que estos actos puedan ser juzgados. Pero esta gran variedad de fuentes probatorias deben tener acceso al proceso judicial a través de alguno de los medios de prueba legalmente previstos.

El advenimiento del fenómeno del “Big Data” ha transformado la tradicional “prueba electrónica”, por cuanto surge la necesidad de un análisis holístico de las múltiples interacciones entre todos aquellos sistemas de información que contienen datos relevantes para el esclarecimiento de los hechos.

Entonces, no solo es necesario comprender cabalmente cómo juegan los principios probatorios en materia de recolección de prueba informática, determinar cómo funcionarán los medios de prueba clásicos en un contexto informático y analizar la potencialidad acreditativa de los nuevos tipos de datos surgidos a partir de la utilización de las nuevas tecnologías, sino que es preciso además, que se estudie el fenómeno informático del Big Data, en lo que hace a la gestión de la información y al cúmulo de datos que se producen.

En consecuencia surgen las siguientes preguntas: ¿ Cuáles son las claves para la incorporación de la prueba digital en un juicio, cuando se trata de Big Data?; ¿Son las

correlaciones más importantes que la causalidad?; ¿Cómo se distingue el Big Data en el contexto de la evidencia digital dentro del proceso de investigación judicial?; y siguiendo de cerca las políticas y legislación vigente ¿Cuáles son las condiciones y estrategias del uso del paradigma del Big Data en el contexto de la evidencia digital?.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación del presente trabajo tiene como eje central el análisis de las condiciones y estrategias del uso del Big Data desde la perspectiva de sus implicaciones jurídicas en el contexto de la evidencia digital, siguiendo de cerca las políticas y legislación vigentes.

Las evidencia digital es un tema que está avanzando en diversas latitudes dentro nuestro territorio nacional, respetando la fisonomía, características y tiempos de cada administración de justicia.

En la provincia de Santiago del Estero en particular, existen diversos avances en proceso de concreción, los que se pueden citar:

- "*Protocolo de actuación para la extracción de evidencia digital y su vinculación con los códigos de procedimientos, de Santiago del Estero, en materia de prueba científica, en un gabinete de investigación forense*". Proyecto del año 2017 de la Facultad de Ciencias para la Innovación y el Desarrollo de la Universidad Católica de Santiago del Estero.
- "*Articulación entre la Evidencia Digital y el Código de Procedimiento, en Materia de Prueba Científica*". Paper presentado en el Congreso Iberoamericano de Investigadores y Docentes de Derecho e Informática CIIDDI 2018.
- "*Creación del Gabinete de Informática Forense (GIF), del Poder Judicial de Santiago del Estero*" [7].
- Proyecto de Investigación de Cátedra: "*Aproximación teórica de las estrategias de delivery de datos unificados del ámbito organizacional*, 2017*,* de la Facultad de Ciencias para la Innovación y el Desarrollo de la Universidad Católica de Santiago del Estero.
- Proyecto de Investigación de Cátedra: "*Big Data: Modelado y visualización de grandes volúmenes de datos jurídicos"*, 2018, de la Facultad de Ciencias para la Innovación y el Desarrollo de la Universidad Católica de Santiago del Estero.

Se trabajará en formular una Guía de Referencia con el fin de mejorar el tratamiento del Big Data, en el marco de la evidencia digital, siguiendo de cerca las políticas y legislación vigentes.

3. RESULTADOS ESPERADOS

El objetivo de esta línea de investigación plantea *establecer una Guía de Referencia de las características diferenciales del Big Data como fuente de datos en el contexto de la evidencia digital y su incidencia en el proceso de investigación judicial, siguiendo de cerca las políticas y legislación vigentes*.

Para lograr ese objetivo, este trabajo se centrará en los objetivos específicos:

- Relevar y examinar los diferentes enfoques técnicos y estratégicos del Big Data disponibles en la actualidad.
- Analizar el contexto del panorama actual de los delitos informáticos basados en big data.
- Distinguir las condiciones del fenómeno de Big Data como fuente del delito o como producto del análisis.
- Identificar formas, que desde un Procedimiento, mejoren el tratamiento de la evidencia digital desde su hallazgo hasta su presentación.
- Formular una guía con el fin de mejorar el tratamiento del Big Data, en el marco de la evidencia digital.

Las actividades que se llevarán a cabo son las siguientes:

- Buscar, recopilar, clasificar y analizar

publicaciones, artículos y libros relacionados con la problemática planteada.

- Examinar los diferentes enfoques técnicos y estratégicos del Big Data.
- Estudiar el Big Data en el contexto de la evidencia digital.
- Identificar los puntos de articulación entre la evidencia digital y las políticas y legislación vigentes.
- Distinción de las condiciones del Big Data como fuente de datos o como herramienta predictiva dentro en el proceso de investigación judicial.
- Documentar las condiciones y estrategias del uso del Big Data como fuente de datos en el contexto de la evidencia digital
- Analizar la incidencia del Big Data en el proceso de investigación judicial, a la luz de las políticas y legislación vigentes

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 dos docentes de la carrera de Informática, una con dedicación simple y otra semiexclusiva, y un asesor legal externo.

El grupo hace difusión y formación de recursos humanos desde las asignaturas: Informática, Análisis Numérico e Ingeniería de Software.

Asimismo, se considera de gran interés la incorporación de becarios, para motivar a los alumnos de Ingeniería en Informática a realizar su trabajo final de grado en el área de este proyecto.

5. BIBLIOGRAFIA

[1] MARÍA LAURA SÁNCHEZ PICCARDI, 2017. Tesis de maestría: "Estado del arte de la industria 4.0". Universidad del Norte Santo Tomás de Aquino, Tucumán, Argentina.

[2] ÓSCAR LÓPEZ, HAVER AMAYA, RICARDO LEÓN. (2001) "Informática forense: generalidades, aspectos técnicos y herramientas". Universidad de Los Andes. Bogotá, Colombia.

[3] SONIA PUIG FAURA, 2014. Tesis doctoral: "La prueba electrónica: sus implicaciones en la seguridad de la empresa". Universidad Ramon Llull, Barcelona, España. Disponible en: https://www.tdx.cat/handle/10803/285237

[4] MERCEDES RIVOLTA,2007. "Medios de prueba electrónicos: estado de avance en la legislación argentina". Disponible en: http://ww.saij.gob.ar/doctrina/dacc070049-rivolta-medios_prueba_electronicos_estado.htm

[5] LORENZO COTINO HUESO, 2017. "Big data e inteligencia artificial. Una aproximación a su tratamiento jurídico desde los derechos fundamentales". Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=6066829

[6] MICHAEL G. NOBLETT. (2000) "Recovering and Examining Computer Forensic Evidence". Disponible en: https://www.fbi.gov/about-us/lab/.../fsc/oct2000/computer.htm

[7] Excmo. Superior Tribunal de Justicia de la provincia de Santiago del Estero, 2017. "Anexo 1, Gabinete De Informática Forense", Disponible en: http://www.jussantiago.gov.ar/jusnueva/Novedades/2017-Anexo1GIF-26-10-2017.pdf

Herramientas estadísticas para data mining y modelos dinámicos

Adriana Mallea[1], Myriam Herrera[2], Jorgelina Carrizo[1], Andrea Salas[3] Cecilia Martínez[3], Leonel Ganga[3]

[1]Departamento de Matemática, FFHA, Universidad Nacional de San Juan
[2]Instituto de Informática, FCEFN, Universidad Nacional de San Juan
[3]Departamento de Informática, FCEFN, Universidad Nacional de San Juan
lamallea@ffha.unsj.edu.ar

RESUMEN

En los últimos años surgió el término Big Data, en el contexto de Data Mining, para referirse a conjuntos de datos tan grandes y complejos que se vuelven difíciles de procesar en un tiempo razonable con aplicaciones tradicionales de análisis de datos. En general, los analistas necesitan analizar datos con variabilidad, por ej., datos que resultan de la agregación de registros individuales en grupos de interés, o datos que representan entidades abstractas como especies biológicas, o regiones como un todo. El Análisis Simbólico de Datos, al ofrecer la posibilidad de agregación de datos al nivel de granularidad elegido por el usuario, mientras se mantiene la información sobre la variabilidad intrínseca, puede desempeñar un papel importante en este contexto. Esta metodología resulta particularmente interesante para el estudio de Economía y Gestión, Marketing, Ciencias Sociales, Geografía, estadísticas sobre datos oficiales, así como para Biología y análisis de datos Geológicos.

Por otra parte, existen campos de la actividad humana; tales como economía, meteorología, administración de empresas; donde parte de los problemas y su solución se presentan en un contexto dinámico. Esto requiere analizar las variables relevantes del problema a resolver a partir de datos recogidos secuencialmente a intervalos regulares de tiempo.

El presente proyecto tiene como finalidad investigar sobre las metodologías estadísticas apropiadas para el tratamiento de grandes bases de datos; en particular del Análisis Simbólico de Datos; como así también estudiar modelos aleatorios para el tratamiento de variables indexadas en el tiempo. Es claro que existen problemas provenientes de diversas áreas del conocimiento donde se necesita de ambos tipos de herramientas, por ejemplo cuando aparecen series temporales o de espacio-tiempo que toman valores en intervalos o incluso distribución de valores. Las metodologías se aplicarán a datos reales.

.

Palabras clave: Data Mining, Modelos, Dinámicos

CONTEXTO

El proyecto *Herramientas estadísticas para Data Mining y modelos dinámicos* se encuentra en su segundo año de ejecución. Es un proyecto acreditado, de carácter bi-anual y financiado por la UNSJ. Tiene como unidades ejecutoras en primer lugar el departamento de Matemática de la FFHA y en segundo

lugar el Instituto de Informática de la FCEFN de la UNSJ. La línea de investigación corresponde a minería de datos, en un marco más general de Ciencia de los Datos.

1. INTRODUCCIÓN

En la actualidad, gracias a los avances en tecnología y recursos informáticos, es sencillo almacenar información, lo que lleva a tener grandes bases de datos. Respuestas referentes a cuestiones formuladas sobre estas bases se pueden dar mediante *Data Mining.*

Data Mining [13] es una poderosa tecnología con gran potencial para ayudar a las compañías a concentrarse en la información más importante de sus Bases de Datos. Las herramientas de minería de datos predicen futuras tendencias y comportamientos, permitiendo tomar decisiones proactivas en los negocios y conducidas por un conocimiento acabado de la información. Permiten responder a preguntas de negocios que tradicionalmente consumen demasiado tiempo para ser resueltas y a las que los usuarios de esta información casi no están dispuestos a aceptar. Estas herramientas exploran las bases de datos en busca de patrones ocultos, encontrando información predecible que un experto no puede llegar a encontrar porque se encuentra fuera de sus capacidades debido a la complejidad dada por el volumen de información.

Las técnicas de Data Mining se pueden implementar rápidamente en plataformas ya existentes de software y hardware para acrecentar el valor de las fuentes de información existentes y se pueden integrar con nuevos productos y sistemas.

Data Mining está soportado por tres tecnologías:

- Recolección masiva de datos.
- Potentes computadoras con multiprocesadores.
- Algoritmos de Data Mining.

En los últimos años surgió el término “Big Data”, en el contexto de Data Mining, para referirse a conjuntos de datos tan grandes y complejos que se vuelven difíciles de procesar en un tiempo razonable con aplicaciones tradicionales de análisis de datos. El Análisis Simbólico de Datos, al ofrecer la posibilidad de agregación de datos, desempeña un rol importante, tal como se comentó en el resumen.

Por otra parte, al analizar datos, éstos pueden aparecer secuencialmente, lo que lleva a considerar variables que dependen del tiempo y, por ende, surge la necesidad de trabajar con modelos dinámicos. En este tipo de situaciones, las propiedades evolutivas que muestran las distintas variables son muy diferentes y, en consecuencia, son también muy distintos los modelos que pueden determinar estas variables en función de variables explicativas.

El conocimiento y caracterización diferenciada de la evolución de las diversas variables relevantes, así como el conocimiento y asimilación de los modelos (probabilísticos) que las pueden explicar, constituyen información imprescindible a la hora de tomar decisiones.

Lo antes mencionado justifica la necesidad de investigar sobre las metodologías estadísticas apropiadas para el tratamiento de grandes bases de datos (Big Data); en particular del Análisis

Simbólico de Datos; como así también estudiar modelos de naturaleza estocástica. Existen problemas provenientes de diversas áreas del conocimiento donde se necesita de ambos tipos de herramientas, por ejemplo evolución de datos financieros, de variables meteorológicas, evolución de indicadores socio- demográficos, es decir variables que toman valores en intervalos o incluso distribución de valores.

1. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La ciencia de datos, considerada como una ciencia en sí misma, es en términos generales, la extracción de conocimiento de los datos [16]. Data Mining es una poderosa tecnología con gran potencial para extraer tal conocimiento. Sin embargo, desde el punto de vista estadístico, sus herramientas sólo han sido desarrolladas para trabajar con matrices de datos clásicas, es decir, donde cada unidad es individual y las variables toman un único valor para cada individuo.
El análisis de datos simbólicos (SDA, por sus siglas en inglés) [9, 10] brinda una nueva forma de pensar en Data Science al extender la entrada estándar a un conjunto de clases de entidades individuales. Por lo tanto, las clases de una población dada se consideran unidades de una población de nivel superior a estudiar. Tales clases a menudo representan las unidades reales de interés. Para tener en cuenta la variabilidad entre los miembros de cada clase, las clases se describen por intervalos, distribuciones, conjunto de categorías o números que a veces se ponderan y similares. De esa manera, obtenemos nuevos tipos de datos, llamados "simbólicos", ya que no se pueden reducir a números sin perder mucha información. El primer paso en SDA es construir la tabla de datos simbólicos donde las filas son clases y las variables pueden tomar valores simbólicos. El segundo paso es estudiar y extraer nuevos conocimientos de estos nuevos tipos de datos mediante al menos una extensión de Estadística Computacional y Data Mining a datos simbólicos.
SDA es un nuevo paradigma que abre un vasto dominio de investigación y aplicaciones al proporcionar resultados complementarios a los métodos clásicos aplicados a los datos estándar. SDA también brinda respuestas a los grandes volúmenes de datos (big data) y datos complejos, ya que los primeros se pueden reducir y resumir por clases y los datos complejos, con múltiples tablas de datos no estructurados y las variables no apareadas se pueden transformar en una tabla de datos estructurada con variables apareadas de valores simbólicos.
En este proyecto trabajamos con ambos enfoques, Data Mining y SDA para la extracción de conocimientos.

1. RESULTADOS OBTENIDOS/ESPERADOS

El presente proyecto tiene como finalidad investigar sobre las metodologías estadísticas apropiadas para el tratamiento de grandes bases de datos; en particular del Análisis Simbólico de Datos; como así también estudiar modelos aleatorios para el tratamiento de variables indexadas en el tiempo y aplicarlas a datos reales.
En el primer año de desarrollo del proyecto se ha trabajado fundamentalmente con el estudio de Series Simbólicas de intervalo y de histograma, como así también de Regresión Lineal Simbólica, en seminarios internos de investigación

desarrollados por los integrantes del proyecto.

Las metodologías propuestas en papers relacionados se han aplicado a datos provenientes de Economía y Finanzas. En particular, para el caso de series simbólicas de intervalo se han usado técnicas de pronóstico de indicadores macroeconómicos. Se han aplicado distintos métodos de regresión simbólica de intervalo y comparado sus performances, usando medidas de error adaptadas al caso simbólico, en un problema de costos.

Por otra parte se han aplicado técnicas del Análisis Simbólico de Datos a datos provenientes de la Encuesta Permanente de Hogares del Gran San Juan, correspondientes al tercer trimestre del año 2016, con el objetivo de caracterizar a éstos hogares y analizar la situación laboral en el Gran San Juan en relación al Nivel de Estudio.

Los resultados obtenidos se han presentado en congresos nacionales e internacionales.

En este segundo año de ejecución se espera profundizar el estudio de herramientas del SDA (Resultados del Workshop Advances in Data Science for Big and complex data, University Paris Duphine, January 2019) y aplicar me todologías de Clustering Simbólico, Regresión Simbólica y Series de Tiempo simbólicas a datos reales o simulados.

1. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por docentes investigadores de dos facultades de la UNSJ, algunos de ellos son jóvenes investigadores. En el marco del proyecto desarrolla su beca de iniciación a la investigación una egresada de Licenciatura en Matemática, que fue alumna adscripta en el proyecto anterior y actualmente cursa su segundo año en la carrera Maestría en Matemática de la Universidad Nacional de San Luis. Entre los integrantes del proyecto hay cuatro maestrandos, que aplicarán en sus trabajos de tesis las herramientas objeto de la presente investigación.

1. BIBLIOGRAFÍA

2. Arroyo, J. (2008) “Métodos de predicción para series temporales de intervalos e histogramas”. Ph. D. Dissertation, Universidad Pontificia Comillas, Madrid.
3. Arroyo, J., R. Espínola, and C. Maté (2008). “Diferent approaches to forecast interval time series: a comparison in finanance.” Computation Statistics and Data Analysis (submitted).
4. Arroyo, J; Gonzales Rivera, G; Maté, C. (2010) “Forecasting with interval and histogram data Some financial applications”. Handbook of empirical economics and finance, 247-280
5. Billard, L., Diday, E. (2007). Symbolic Data Analysis: Conceptual Statistics and Data Mining. Wiley.
6. Brito, P. (2014): "Symbolic Data Analysis: another look at the interaction of Data Mining and Statistics". WIREs Data Mining and Knowledge Discovery, Volume 4,

Issue 4, July/August 2014, 281–295. DOI: 10.1002/widm.1133

7. Brito, P., Duarte Silva, A. P. (2012): "Modelling Interval Data with Normal and Skew-Normal Distributions". Journal of Applied Statistics, Volume 39, Issue 1, 3-20.
8. Brito, P. (2007): "Modelling and Analysing Interval Data". In: "Advances in Data Analysis", Decker, R., Lenz, H.-J. (Eds.), Series "Studies in Classification, Data Analysis and Knowledge Organization", Springer, Berlin, Heidelberg, New-York, 197-208.
9. Brito, P. (2007): "On the Analysis of Symbolic Data". In: "Selected Contributions in Classification and Data Analysis", Brito, P., Bertrand, P.,
10. Diday, E (2016) "Thinking by classes in data Science: the symbolic data analysis paradigm". WIREs Comput Stal,8:172-205,doi: 0.1002/wics.1384.
11. Diday, E. and Noirhomme-Fraiture, M. (2008). Symbolic Data Analysis and the SODAS Software. Wiley.
12. Gallardo, V; Mallea, L. A. (2016). "Una introducción a Datos Simbólicos". Servicio de Publicaciones de la FFHA-UNSJ. ISBN 978-950-605-837-1
13. Han, A., Hong, Y., Lai, K. and Wang, S. (2008). "Interval time series analysis with an application to the Sterling-Dollar exchange rate", Journal of Systems Science and Complexity, 21 (4), 558-573.
14. Han, J. Kamber, M.(2006), Data Mining. Concepts and Techniques, 2.a ed. Morgan Kaufman.
15. Maia, A.L.S., De Carvalho, F.A.T. and Ludermir, T.D. (2008). "Forecasting models for interval-valued time series", Neurocomputing, 71 (16-18), 3344-3352.
16. Mallea, L.A.; Martínez, E.; Salas, A. (2017) "Series Temporales Simbólicas de Intervalo". Servicio de Publicaciones de la FFHA-UNSJ. ISBN 978-950-605-851-7
17. Taylor, J. W. "Exponential smoothing with a damped multiplicative trend. International Journal of Forecasting", vol. 19, páginas 715_725, 2003.
18. Teles, P. y Brito, M. P. (2005) "Modelling interval time series data". En Proceedings of the 3rd IASC World Conference on Computational Statistics & Data Analysis.
19. Teles, P. and Brito, P. (2015)." Modelling Interval Time Series with Space-Times processes", communications in Statistics: Theory and Methods, Volume 44, Issue 17. DOI: 10.1080/03610926.2013.782200

Indexando Bases de Datos no Estructurados

Norma Herrera, Darío Ruano, Paola Azar, Daniel Welch
Departamento de Informática
Universidad Nacional de San Luis, Argentina
{nherrera, dmruano, epazar, dwelch}@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/F814), 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. Describimos a continuación los modelos de bases de datos e índices sobre los que nos estamos trabajando.

Bases de Datos de Texto. Una base de datos de texto es un sistema que mantiene una colección grande de texto y que provee acceso rápido y seguro al mismo. Sin pérdida de generalidad, asumiremos que la base de datos de

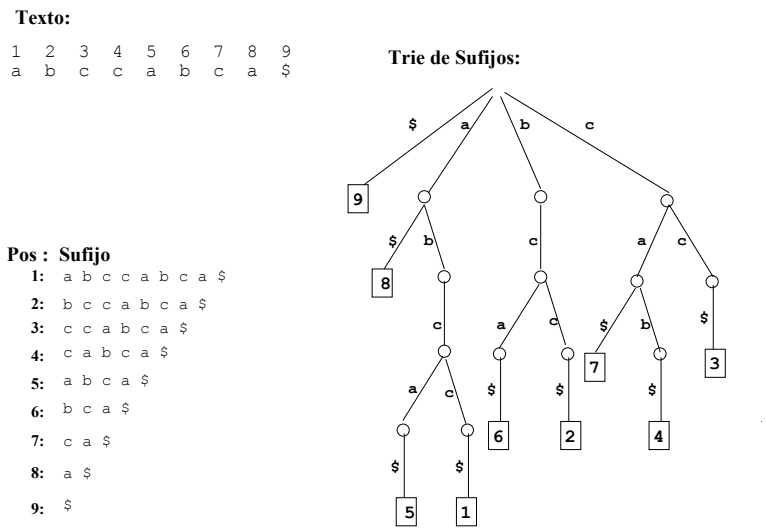


Figure 1: Un ejemplo de un texto, sus sufijos y su correspondiente trie de sufijos.

texto es un único texto T que posiblemente se encuentra almacenado en varios archivos.

Una de las búsquedas más comunes en bases de datos de texto es la *búsqueda de un patrón*: el usuario ingresa un string P (*patrón de búsqueda*) y el sistema retorna las ocurrencias del patrón P en el texto T. Para resolver eficientemente estas búsquedas se hace necesario construir un índice que permita acelerar el proceso de búsqueda.

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 [4, 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 [5, 6]. Por esta razón, el estudio de índices comprimidos y en memoria secundaria para texto es un tema de interés

Dado un texto $T = t_1, \ldots, t_n$ sobre un alfabeto Σ de tamaño σ, donde $t_n = \$ \notin \Sigma$ es un símbolo menor en orden lexicográfico que cualquier otro símbolo de Σ, definimos ***sufijo de T*** a cualquier string de la forma $T_{i,n} = t_i, \ldots, t_n$ con $i = 1..n$. Cada sufijo $T_{i,n}$ se identifica unívocamente por i; llamaremos al valor i ***índice del sufijo*** $T_{i,n}$.

Un ***trie de sufijos***[4] es un *trie* construido sobre el conjunto de todos los sufijos del texto, en el cual cada hoja mantiene el índice del sufijo que esa hoja representa. La figura 1 muestra un ejemplo de un texto, los sufijos del texto con la posición del texto donde cada sufijo comienza y su correspondiente trie de sufijos.

La representación habitual de un trie consiste en mantener en cada nodo los punteros a sus hijos, junto con el rótulo correspondiente a cada uno de ellos [7].

En [9] 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.

Bases de Datos Métrico-Temporales. Este modelo permite almacenar objetos no estructurados con tiempos de vigencia asociados y realizar consultas por similitud y por tiempo

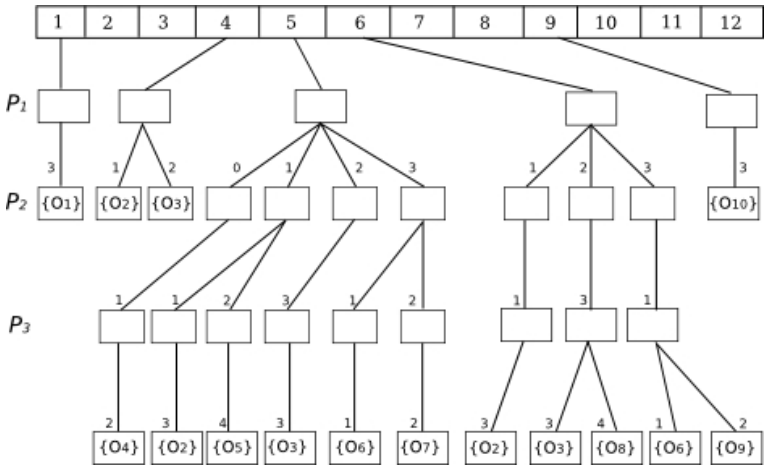


Figure 2: Un ejemplo de un H-FHQT

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 triupla (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: $(q, r, t_{iq}, t_{fq})_d = \{o/(o, t_{io}, t_{fo}) \in X \wedge d(q, o) \leq r \wedge (t_{io} \leq t_{fq}) \wedge (t_{iq} \leq t_{fo})\}$ Esta consulta implica buscar todos los objetos o de la base de datos que estén a una distancia a lo más r de q, y cuyo tiempo asociado t se solapa con el 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. El *Historical-FHQT*(H-FHQT) [3] es un índice métrico-temporal que utiliza tanto la componente métrica como la temporal para resolver eficientemente búsquedas métrico-temporales. Este índice consiste en una lista de instantes válidos donde cada uno contiene un FHQT que indexa a todos los objetos vigentes en dicho instante. Los FHQT tienen distintas profundidades, es decir distintas cantidades de pivotes, en función de la cantidad de elementos que deban indexar.

Si bien en un H-FHQT la cantidad de pivotes en distintos instantes de tiempo varía, siempre se trabaja sobre el mismo conjunto base de pivotes; esto significa que si en el instante i necesito k_i pivotes y en el instante j necesito k_j pivotes con $k_i < k_j$, entonces los primeros k_i pivotes de ambos instantes son iguales.

La figura 2 muestra un ejemplo de un H-FHQT construido sobre el intervalo de tiempo [1..12]. Se puede observar que el objeto o_2 está vigente en el intervalo de tiempo [4..6] y se encuentra a distancia 1 de p_1, a distancia 1 de p_2 y a distancia 3 de p_3.

3 Líneas de Investigación

3.1 Bases de Datos de Texto

Como mencionáramos en la sección anterior, en [9] 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.

Hemos diseñado e implementado una técnica de paginación para este índice basándonos en la representación secuencial del mismo. Nuestra técnica de paginación consiste en una extensión para árboles r-arios del paginado utilizado en el Compact Pat Tree [2]. Al igual que el algoritmo presentado en [2], 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:

• Se coloca 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 ocurre 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.

• Se aplica 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.

Los resultados obtenidos con esta primer técnica nos ha permitido analizar el funcionamiento de la misma y nos encontramos diseñando mejoras a la misma.

Paralelemamente estamos trabajando sobre técnicas de compresión para el trie de sufijos, con el obejtivo final de implementar una versión compacta y en memoria secundaria de este índice.

3.2 Modelo Métrico-Temporal

En este ámbito nuestro obejtivo es el diseño de índices en memoria secundaria. Específicamente hemos estamos trabajando con el H-FHQT. 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:

1a: 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 de disco.

1b: 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 nia cada uno de los árboles FHQT correspondientes a cada instante de tiempo entran 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.

Nos encontramos implementando esta primer versión de paginado del H-FHQT.

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 bases de datos métrico temporales. 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 ámbito de 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] D. Clark and I. Munro. Efficient suffix tree on secondary storage. In *Proc. 7th ACM-SIAM Symposium on Discrete Algorithms*, pages 383–391, 1996.

[3] 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, Agentina, 2007.

[4] 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.

[5] 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.

[6] 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.

[7] A. Thomo M. Barsky *, U. Stege. A survey of practical algorithms for suffix tree construction in external memory. In *Software: Practice and Experience*, 2010.

[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] 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.

Introducción a las Bases de Datos de Grafos: Experiencias en Neo4j

Lic. Cristina Vera - Lic. Silvina Migani

Departamento de Informática / Facultad de Ciencias Exactas, Físicas y Naturales / UNSJ
Av. Ignacio de la Roza 590 (O)
Teléfonos: 4260353 - 4260355
civera2@yahoo.com.ar; silvina.migani@gmail.com

Resumen
Las bases de datos NoSQL surgieron como una alternativa de solución a problemas no resueltos eficientemente por las tradicionales bases de datos relacionales. Sin embargo, el término NoSQL abarca diferentes tipos de bases de datos, cada uno con sus características intrínsecas que le confieren un comportamiento más o menos apropiados para diferentes escenarios de aplicación. Las bases de datos de grafos se encuentran dentro de esa gran y variada familia de bases de datos.
Esta línea de investigación surge como una iniciativa de las dos docentes que constituyen el equipo de cátedra de las asignaturas de bases de datos del Departamento de Informática, con la finalidad de profundizar en el estudio de esta nueva y prometedora generación de bases de datos.

Palabras Clave: Bases de Datos – Bases de Datos NoSQL - Bases de Datos de Grafos

Contexto
Las cátedras Bases de Datos y Tópicos Avanzados de Bases de Datos de Departamento

de Informática de la FCEFyN de la UNSJ constituyen el contexto dentro del cual se desarrollan las actividades de investigación.

1. Introducción

El uso masivo de Internet generó nuevas formas de producir y compartir información. La Web 2.0 es una web dinámica y participativa, donde los usuarios son protagonistas activos, generando y compartiendo contenidos, opinando y participando. Esta situación provoca la generación de una enorme cantidad de datos altamente relacionada, que necesita ser almacenada y manipulada eficientemente, con un elevado grado de disponibilidad y que además, comúnmente no se ajusta a una estructura rígida. En este escenario, surgieron las bases de datos NoSQL, ya que las bases de datos relacionales no pudieron satisfacer adecuadamente esas exigencias. Las bases de datos de grafos constituyen una de las alternativas dentro de la gran y variada familia de bases de datos NoSQL. Así, esta propuesta pretende profundizar en el estudio de bases de datos y de gestores de bases de datos basados en grafos (SGBDGs), con el propósito de identificar, distinguir, experimentar y valorar sus características específicas.
Este artículo describe brevemente los temas abordados hasta el momento dentro de la temática planteada:

Sistemas de Bases de Datos basados en Grafos (SBDGs)

Antes de caracterizar a los SBDGs, se presenta la definición matemática de grafo. Un grafo simple *G* está definido por *V(G),* un conjunto finito y no vacío de elementos llamados vértices o nodos, y por *E(G),* un conjunto finito de pares no ordenados de elementos distintos de *V(G)* llamados arcos o aristas. Una arista *{v, w}* une los vértices *v* y *w*, comúnmente denotada como *vw* [1].

Un SGBDG, es un sistema de base de datos específicamente diseñado para poseer las siguientes capacidades [2]:

- Administrar datos de tipo grafo. Es decir, su modelo de datos lógico está basado en alguna de las variantes de la definición matemática básica de grafo, como por ejemplo, grafos dirigidos o no dirigidos, con vértices y arcos etiquetados o no etiquetados, con propiedades en nodos y arcos, hipergrafos e hipernodos [3][4]. Consecuentemente, las operaciones CRUD (Create-Read-Update-Delete, operaciones básicas de los SGBDs) trabajan sobre grafos y sus elementos.

  Sin embargo esto no significa que en el almacenamiento subyacente efectivamente se encuentren grafos. Algunos SGBDGs utilizan almacenamiento nativo [5]; es decir, sus estructuras de datos físicas están diseñadas y optimizadas para almacenar y administrar grafos. También existen gestores que mapean los grafos a otras estructuras [5], como por ejemplo tablas u objetos. En cuanto a la implementación de las conexiones entre nodos, también se presentan diferentes enfoques, algunos que eligen la adyacencia sin índices, y otros que no.
- Satisfacer los principios básicos de todo SGBD, es decir, el almacenamiento persistente, la independencia física/lógica, la integridad y la consistencia de los datos.

Modelo de Grafos vs Modelo Relacional

El modelo de datos relacional y los SGBDs relacionales son ampliamente conocidos y populares. Por ello, a continuación, en la Tabla 1, se presentan algunos de los elementos básicos del enfoque relacional y sus análogos en el modelo de grafos. Cabe aclarar que dicha correspondencia no es absoluta ya que cada pareja de términos presenta particularidades significativas.

Bases de Datos Relacionales	Bases de Datos de Grafos
Filas	Nodos
Columnas	Propiedades
Nombre de las Tablas	Etiquetas en Nodos/Aristas
Claves Foráneas	Aristas entre Nodos

Tabla 1:
Comparación Modelo Relacional – Modelo de Grafos

Algunos SGBDGs presentes en el mercado

Hoy en día existen numerosos SGBDs basados en grafos. A continuación se mencionan en orden de popularidad según [6] algunos: Neo4j, Microsoft Azure Cosmos DB, OrientDB, ArangoDB, Virtuoso y JanusGraph.

Ensayos Realizados

Para poder experimentar las características inherentes a este tipo de sistemas de bases de datos, se siguieron los siguientes pasos:

1. Definición de un escenario sencillo apropiado para ser resuelto con este tipo de tecnología. Así, se concibió un sistema recomendador que brindara información relativa a las sucursales de San Juan Servicio cuando una persona tiene que efectuar un pago, en base a la opinión de sus amigos. Para ello, se registraron las experiencias de pago, ponderadas a través de una calificación (valor numérico entero entre 1 y 10).

2. Elección de un SGBDG donde concretar las experiencias. Se decidió usar Neo4j, sigue el modelo de grafo de propiedades y es una herramienta muy útil para prototipado rápido y para hacer pruebas de concepto, ya que es sencilla y no presenta complicaciones en la configuración [7]. Además, provee el lenguaje Cypher que permite expresar las operaciones CRUD de manera bastante natural [8].

3. Implementación del ejemplo planteado. Luego de instalar Neo4j Desktop 1.1.15 se generó el grafo correspondiente a través de operaciones escritas en el lenguaje Cypher.

Algunas de las operaciones de creación de nodos y aristas se muestran a continuación:

CREATE (Maria:Persona {nombre:'María', fechanac:'26-01-1990'})

CREATE (Suc1:Sucursal {nombre:'Casa Matriz', localidad:'San Juan', Direccion: 'Scalabrini Ortiz 1285 N', horaten:'8-13'})

CREATE (Maria)-[:Paga {califvisita:6}] ->(Suc1)

CREATE (Maria)-[:Amigo]->(Juan)

Analizando la relación Paga, se observa que contiene la propiedad calificación. En el ejemplo, María calificó a la Casa Matriz (Suc1) con 6.

El grafo resultante luego de la ejecución de las sentencias de creación de todos los nodos y aristas se muestra a continuación en la Figura 1.

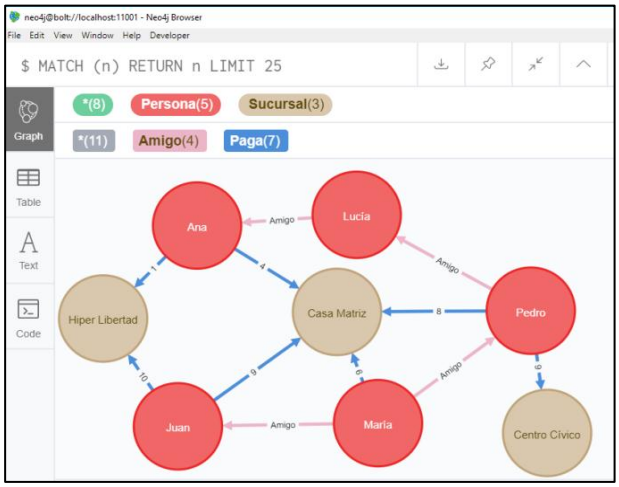


Figura 1: Grafo ejemplo

Asimismo, a modo de ejemplo, se muestran a continuación algunas de las consultas implementadas para el sistema recomendador propuesto.

C1: Obtener para cada sucursal, su calificación promedio.

MATCH (S:Sucursal)-[Pa:Paga]-() WITH S, AVG(Pa.califvisita) AS prom
return S.nombre,prom

La Figura 2 muestra el resultado correspondiente[1].

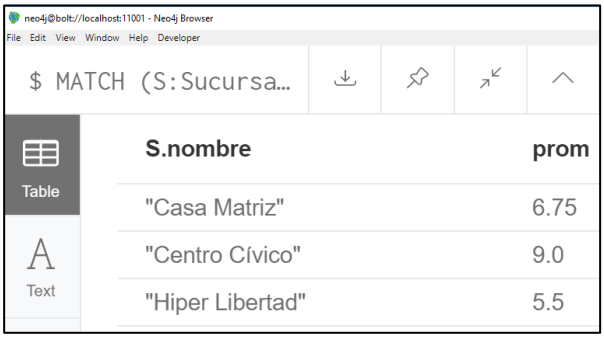


Figura 2: Resultado consulta C1

[1] Esta consulta no devuelve un grafo, por eso el resultado se muestra como una tabla.

C2: Obtener la sucursal con máxima calificación promedio.

MATCH (s:Sucursal)-[Pa:Paga]-() WITH s, AVG(Pa.califvisita) AS prom
RETURN s ORDER BY prom DESC
LIMIT 1

El resultado obtenido se presenta en la Figura 3.

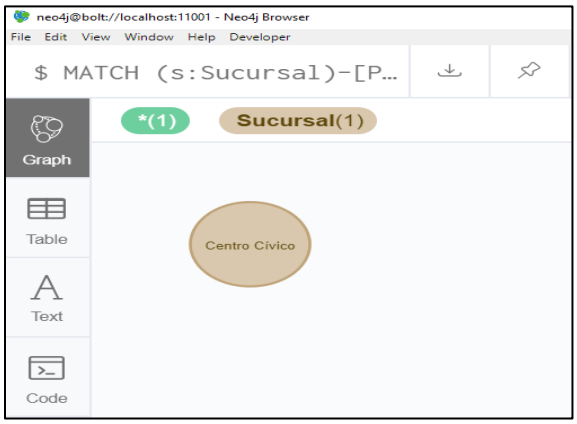


Figura 3: Resultado de la consulta C2

C3: Considerando las calificaciones de los amigos de Juan, obtener la sucursal con mayor calificación promedio.

MATCH (p:Persona{nombre:'Juan'})-[A:Amigo]-(p2:Persona)-[pa:Paga]-> (s:Sucursal) WITH s, AVG(pa.califvisita) AS prom
RETURN s as Sucursal,prom as Promedio
ORDER BY prom DESC
LIMIT 1

En la Figura 4 se puede observar el resultado.

Figura 4: Resultado de la consulta C3

C4: Considerando las calificaciones de los amigos de Juan y de los amigos de los amigos de Juan, obtener la sucursal con mayor calificación promedio.

*MATCH (p:Persona{nombre:'Juan'})-[A:Amigo*1..2]-(p2:Persona)-[pa:Paga]-> (s:Sucursal) WITH s, AVG(pa.califvisita) AS prom*
RETURN s as Sucursal,prom as Promedio
ORDER BY prom DESC
LIMIT 1

El resultado obtenido se presenta en la Figura 5.

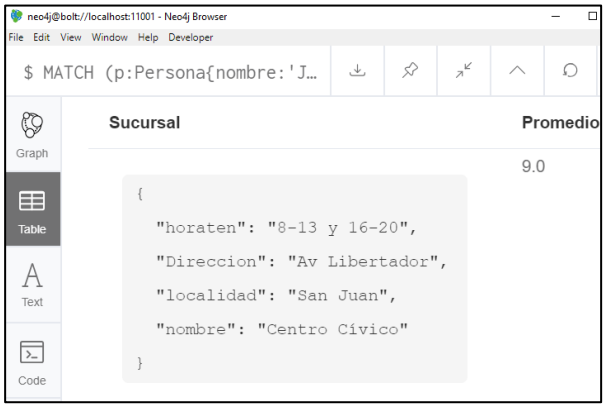


Figura 5: Ejecución de la consulta C4

Como se puede observar, la sucursal recomendada a Juan al considerar los pagos realizados por sus amigos (C3) no es la misma que la obtenida al considerar además de los pagos de los amigos de Juan, los pagos de los amigos de los amigos de Juan (C4); ya que se consideran instancias diferentes de la relación Paga.

En base a los ejemplos planteados e implementados, se pudo distinguir y valorar el potencial de esta tecnología, fundamentalmente en lo que concierne a:

- La gran flexibilidad en la estructura de los datos. Tanto los nodos como las aristas no necesitan ajustarse a una estructura predeterminada.

- El hecho de contar con un lenguaje de consulta declarativo potente e intuitivo, que hace muy sencilla la tarea de codificar consultas.
- La posibilidad de navegar eficientemente a través de relaciones complejas, incluso recursivas (relaciones entre nodos del mismo "tipo") a cualquier nivel de profundidad, incluso no fijo. Capacidad no ofrecida convenientemente por los sistemas SQL.

Líneas de Investigación y desarrollo

La investigación tiene como eje central el estudio y la experimentación de bases de datos NoSQL basadas en grafos.

Resultados y Objetivos

Los propósitos fundamentales establecidos fueron:

- Profundizar en el estudio de las bases de datos de grafos en el marco de los SGBD NoSQL.
- Investigar y experimentar SGBDs dentro de esta tipología.
- Asesorar trabajos finales en el área.
- Elaborar material bibliográfico.
- Fortalecer las cátedras de Bases de Datos del Departamento de Informática.

Sin embargo, dado que el trabajo de investigación se encuentra en sus inicios, los resultados obtenidos hasta el momento son:

- Se han estudiado las características intrínsecas del modelo de datos de grafos, identificando clasificaciones dentro de este grupo.
- Se han investigado gestores de bases de datos de grafos presentes en el mercado.
- Se ha profundizado en el estudio de Neo4j Desktop 1.1.15 como SGBD de experimentación.
- Se han definido e implementado algunos ejemplos concretos en Neo4j.
- Se está desarrollando una tesis de grado y una de maestría sobre la temática NoSQL, que incluyen los SGBDGs.

Formación de Recursos Humanos

- Se está desarrollando una tesis de grado y una de maestría sobre la temática NoSQL, que incluyen los SGBDGs.

Bibliografía

[1] Wilson, R. J., Wilson, R. J., Wilson, R. J., & Wilson, R. J. (1972). Introduction to graph theory (Vol. 107). London: Longman.

[2] Angles, R., & Gutierrez, C. (2018). An introduction to Graph Data Management. In Graph Data Management (pp. 1-32). Springer, Cham.

[3] Sasaki B., Chao J. & Howard R. (2018). Graph Databases for Beginners.

[4] Introduction to Graph Databases.

Disponible en https://neo4j.com/graphacademy/online-training/introduction-to-neo4j/part-1/. Fecha última visita Marzo 2019.

[5] DB-Engines Ranking of Graph DBMS. Disponible en https://db-engines.com/en/ranking/graph+dbms. Fecha última visita Marzo 2019.

[6] Robinson, I., Webber, J., & Eifrem, E. (2013). Graph databases. " O'Reilly Media, Inc.".

[7] Sitio oficial de Neo4j. Disponible en https://neo4j.com Fecha última visita Marzo 2019.

[8] The Neo4j Cypher Manual v3.5. Disponible en https://neo4j.com/docs/cypher-manual/3.5/. Fecha última visita Marzo 2019.

MÉTODOS DE ACCESO MÉTRICO-ESPACIALES

Andrés Pascal, Anabella De Battista
Grupo de Investigación en Bases de Datos (GIBD)
Facultad Regional Concepción del Uruguay
Universidad Tecnológica Nacional
andrespascal2003@yahoo.com.ar, anadebattista@gmail.com

Norma Herrera
Departamento de Informática
Universidad Nacional de San Luis
nherrera@unsl.edu.ar

RESUMEN

Tradicionalmente, los datos que contienen las bases de datos están estructurados en tuplas y son comparables a través de operadores relacionales. Para acelerar este tipo de consultas existen índices eficientes, tales como B+-Tree. Sin embargo, cada vez es más importante el almacenamiento de objetos no estructurados, que no se pueden comparar por igualdad, para los cuales dichos índices no son aplicables. Algunos ejemplos son: imágenes (rostros, radiografías, pinturas, marcas, paisajes, etc.), texto plano y semiestructurado (documentos, archivos XML, etc.), sonidos (música, voz, etc.) y objetos espaciales (ciudades, rutas, puntos de interés, etc.). Ante esta situación, han surgido otras formas de consultas, siendo algunas de las más importantes las espaciales y las por similitud.

Un aspecto no estudiado aún, es la combinación de estos dos tipos de búsqueda, e.g. "encontrar objetos similares a uno dado, ubicados dentro de un área". Estos tipos de consultas son importantes en especial en los Sistemas de Información Geográfica y aún no existen métodos de acceso que los soporten.

En este proyecto estudiamos distintos aspectos referidos al procesamiento de consultas métrico-espaciales, las funciones de distancia a utilizar, y el uso de paralelismo en GPU para hacer más eficiente el procesamiento de las mismas.

Palabras clave: consultas por similitud, espacios métricos, consultas espaciales, índices

CONTEXTO

Este trabajo se inscribe dentro del Proyecto homologado "Consultas por Similitud y Espaciales en Bases de Datos de Objetos No Estructurados" (UTI3842TC) cuyo objetivo es el desarrollo de métodos y técnicas que mejoren la eficiencia y efectividad de los métodos de búsqueda de objetos no estructurados.

Como parte del proyecto se firmó un convenio con la Municipalidad de Urdinarrain para realizar el relevamiento y geocodificación de los Comercios, Industrias y Estudios Profesionales de dicha ciudad y para el desarrollo de una aplicación que los registre y permita consultas para la toma de decisiones.

Se establecieron comunicaciones con las Oficinas de Marcas y Señales de las provincias de Entre Ríos y Buenos Aires para el desarrollo de una aplicación de búsqueda por similitud de Marcas de Ganado.

Este proyecto es continuación de los proyectos "Procesamiento eficiente de consultas en nuevos modelos de bases de datos" (Incentivos: 25-D059) y "Métodos de Acceso, Consultas y Aplicaciones en Modelos de Bases de Datos no Convencionales" (Incentivos: 25-D040).

1. INTRODUCCIÓN

En las grandes bases de datos, para realizar búsquedas con eficiencia se requiere de algún soporte y organización especial a nivel físico. Las bases de datos clásicas y en particular las relacionales, organizan los datos en conjuntos de registros de tamaño fijo que contienen campos completamente comparables. Esto les permite realizar consultas exactas o por prefijo con costo menor a O(n), mediante estructuras de datos auxiliares llamadas índices, tales como el

B+-Tree o las Hash-tables. Esta manera de organizar los datos y consultar no es adecuada cuando los objetos en cuestión son no estructurados, para los cuales no tiene sentido las comparaciones por igualdad, y donde las consultas incorporan otras dimensiones como el espacio o el tiempo. Por ejemplo, carece de sentido buscar un rostro exactamente igual, píxel a píxel, a algún otro contenido en una base de datos de rostros de personas, o consultar por igualdad una polilínea que representa una ruta dentro de una base de datos espacial.

En este contexto, surgen como respuesta al requerimiento de almacenar y consultar objetos no estructurados y con algún aspecto espacial o temporal, los modelos de Bases de Datos Espaciales, Temporales, Espacio-Temporales, los Espacios Métricos y las Bases de Datos Métrico-Temporales.

Las Bases de Datos Espaciales [1, 2, 3, 4, 5, 6, 7] permiten procesar objetos con alguna referencia espacial y que poseen normalmente una estructura compleja. Un dato espacial se representa usualmente a través de un punto, una polilínea o un polígono. La recuperación y actualización de estos tipos de datos espaciales se basan no sólo en el valor de ciertos atributos, sino también en la ubicación espacial del objeto. Por ejemplo, nos podría interesar obtener los terrenos geográficamente adyacentes a uno dado, o encontrar todas las estaciones de servicio al lado de una ruta. Las bases de datos espaciales se utilizan en muchas áreas, destacándose en particular los Sistemas de Información Geográfica (SIG) [8, 9, 10]. Un SIG es principalmente una herramienta que permite capturar, almacenar, manipular, analizar y mostrar información geográficamente referenciada con el objetivo de resolver problemas complejos de planificación y gestión. En una sociedad donde la información y la tecnología son dos pilares fundamentales, los SIG proveen el marco tecnológico adecuado para el manejo de información geográfica y permiten canalizar la gestión de todo aquello que presente una componente geográfica susceptible de ser aprovechada.

Por otro lado, cuando se almacenan objetos complejos con estructuras variables tales como las imágenes, las consultas que tienen sentido son las llamadas Consultas por Similitud [12, 13, 14]. Una forma de modelar este tipo de bases de datos es mediante Espacios Métricos [11]. Un espacio métrico es un par *(U, d)* donde U es un universo de objetos y *d* es la función de distancia definida entre los elementos de *U*, que mide el grado de similitud (disimilitud, estrictamente hablando) entre ellos y que posee ciertas propiedades que la hacen "métrica".

Una de las consultas típicas en este modelo de bases de datos es la búsqueda por rango: dado un elemento *q* incluido en *U* al que llamaremos *query* y un radio de tolerancia *r*, una búsqueda por rango consiste en recuperar los objetos de la base de datos que se encuentren como máximo a distancia *r* de *q*. Para resolver estas consultas sin realizar O(n) evaluaciones de distancias se utilizan índices que permiten ahorrar cálculos durante el proceso de búsqueda.

En algunos casos, cada vez más frecuentes, las funciones de distancia no cumplen alguna de las propiedades de una función métrica [15, 16]. En especial si la función no respeta la desigualdad triangular, los índices métricos no pueden ser utilizados.

Estos temas constituyen un área de investigación abierta y de gran importancia dado que se basan en necesidades de aplicación reales que aún no están resueltas. Si bien algunos motores de bases de datos como Postgres, Oracle o Informix, ya han incorporado el aspecto temporal o espacial, aún no permiten manejar o tienen limitaciones importantes en cuanto a los demás modelos.

La paralelización de problemas de recuperación de información sobre arquitecturas multi-core, ha sido estudiada desde diversas perspectivas [17, 18, 19, 20, 21, 22, 23]. Por ejemplo, el algoritmo de consulta NN_k (k-Nearest Neighbors) que se utiliza en muchos sistema de recuperación de datos ya ha sido paralelizado con buenos resultados [24, 25].

Actualmente una comunidad importante de investigadores se encuentra dedicada al estudio de la aceleración de diversos algoritmos a través del uso de Unidades de Procesamiento Gráfico (GPU: Graphics Processing Unit). Estas soluciones explotan la naturaleza de las GPUs para alcanzar resultados significativamente más eficientes que los obtenidos mediante el uso

normal de una CPU, utilizando hardware de bajo costo y amplia disponibilidad.
En este proyecto profundizamos el estudio de los distintos aspectos involucrados al procesamiento de consultas por similitud y espacial al mismo tiempo, para las cuales no existen métodos de acceso específicos. En segundo lugar, estudiamos las búsquedas por similitud cuando las funciones no son métricas, incluyendo el desarrollo de estructuras de acceso para estos casos. Y por último, el uso de GPUs para acelerar los mecanismos de búsqueda a través del uso de paralelismo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Actualmente las líneas de investigación en las cuales se estás desarrollando las actividades del proyecto son:

- Métodos de Acceso Métrico-Espaciales
- Funciones de Distancia y Extracción de Características de Objetos no Estructurados
- Paralelismo sobre GPU

3. RESULTADOS OBTENIDOS/ESPERADOS

En proyectos anteriores se definió el modelo métrico-temporal, se desarrollaron métodos de acceso para procesar consultas que combinan similitud y tiempo, se definieron tipos de consultas sobre este modelo y se construyeron aplicaciones que las soportan.
El objetivo principal de este proyecto es el estudio de nuevas tecnologías para la resolución de Consultas por Similitud y Espaciales, incluyendo:
- Definición del modelo métrico-espacial, incluyendo consultas, métodos de acceso y aplicaciones para este nuevo modelo.
- Estudio y desarrollo de funciones de distancia no-métricas
- Diseño de Métodos de Acceso no-métricos
- Uso de paralelismo (GPUs) para mejorar la eficiencia del procesamiento de consultas métrico-temporales y no-métricas.

Actualmente se han desarrollado funciones de distancia métricas y no-métricas para medir la similitud de imágenes por forma [26, 27], y se ha paralelizado la extracción de características utilizando GPU-Cuda, obteniendo buenos resultados.
También se ha diseñado un primer método de acceso métrico espacial, y se están realizando ajustes y validaciones para realizar los experimentos correspondientes para medir su eficiencia.

4. FORMACIÓN DE RECURSOS HUMANOS

Participan en este proyecto, además de los investigadores a cargo, dos tesistas de la Maestría en Ciencias de la Computación, dos graduados (uno de ellos con beca BINID) de la carrera Ingeniería en Sistemas de Información y todos los años se suman al menos dos alumnos becarios.

5. BIBLIOGRAFÍA

[1] Gandhi, V., Kang, J. M., Shekhar, S.: Spatial Databases, Encyclopedia of Computer Science and Engineering, Wiley, Cassie Craig (Eds.), (2009).

[2] Gaede, V., Günther, O.: Multidimensional access methods. ACM Comput. Surv. 30, 2 (1998).

[3] Baeza-Yates, R., Ribeiro-Neto: Modern Information Retrieval. Addison Wesley (1999).

[4] Beckmann, N., Kriegel, H., Schneider, R., Seeger, B.: The R*-tree: an efficient and robust access method for points and rectangles. ACM International Conference on Magamenet of Data (1990).

[5] Bereczky, N., Duch, A., Németh, K., Roura, S.: Quad-kd trees: A general framework for kd trees and quad trees. Theoretical Computer Science, Volume 616, 22 (2016) ISSN 0304-3975.

[6] Felipe, I., Hristidis, V., Rishe, N.: Keyword search on spatial databases. In ICDE'08, pages 656–665 (2008).

[7] Rouquier, J., Alvarez, I., Reuillon, R., Wuillemin, P.: A kd-tree algorithm to discover the boundary of a black box hypervolume. Annals of Mathematics and Artificial Intelligence (2015).

[8] Chen, Y., Suel, T., Markowetz, A.: Efficient Query Processing in Geographic Web Search Engines. In SIGMOD'06 (2006).

[9] Li, Z., Wang, C., Xie, X., Wang, X., Ma, W.: Indexing implicit locations for geographical information retrieval. In GIR'06 (2006).

[10] Elariss, H., Khaddaj, S.:A time cost optimization for similar scenarios mobile GIS queries. Journal of Visual Languages & Computing (2012).

[11] E. Chávez, G. Navarro, R. Baeza-Yates, and J.L. Marroquín. Searching in Metric Spaces. ACM Computing Surveys, 33(3):273–321, September (2001).

[12] 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).

[13] R. Baeza-Yates. Searching: an algorithmic tour. In A. Kent and J. Williams, editors, Encyclopedia of Computer Science and Technology, volume 37, pages 331-359. Marcel Dekker Inc., (1997).

[14] R. Baeza-Yates and G. Navarro. Fast approximate string matching in a dictionary. In Porc. 5th South American Symposium on String Processing and Information Retrieval (SPIRE'98) (1998).

[15] Scheirer, W., Wilber, M., Eckmann, M., Boult, T.: Good recognition is non-metric. Pattern Recognition, Volume 47 (2014).

[16] Chen, S., Ma, B., Zhang, K.: On the similarity metric and the distance metric. Theoretical Computer Science, Volume 410, Issues 24–25 (2009).

[17] Vinkler, M., Havran, V., Bittner, J.: Performance Comparison of Bounding Volume Hierarchies and Kd-Trees for GPU Ray Tracing. Computer Graphics Forum (2015).

[18] Sun, C., Agrawal, D., Abbadi, A.: Hardware acceleration for spatial selections and joins. In: Proc. ACM Intl. Conf. On Managment of Data. (2003) 455–466.

[19] Bandi, N., Sun, C., Abbadi, A., Agrawal, D.: Hardware acceleration in commercial databases: A case study of spatial operations. In: Proc. Intl. Conf. on Very Large Databases, Morgan Kaufmann (2004) 1021–1032.

[20] Govindaraju, N., Lloyd, B., Wang, W., Lin, M., Manocha, D.: Fast computation of database operations using graphics processors. In: Proc. ACM Intl. Conf. On Managment of Data. (2004) 215–226.

[21] Owens, J., Luebke, D., Govindaraju, N., Harris, M., Krüger, J., Lefohn, A., Purcell, T.: A survey of general-purpose computation on graphics hardware. Proc. ACM (2007).

[22] Fatahalian, K., Sugermanm J., Hanrahan, P.: Understanding the Efficiency of GPU Algorithms for Matrix-Matrix Multiplication. Proc. of the ACM SIGGRAPH / EUROGRAPHICS conference on Graphics hardware (2004).

[23] Matsumoto, T., Yiu, M.: Accelerating Exact Similarity Search on CPU-GPU Systems. ICDM, IEEE International Conference on Data Mining (2015).

[24] Cayton, L.: A nearest neighbor data structure for graphics hardware. First International Workshop on Accelerating Data Management Systems Using Modern Processor and Storage Architectures (2010).

[25] Bustos, B., Deussen, O., Hiller, S., Keim, D.: A graphics hardware accelerated algorithm for nearest neighbor search. Computational Science ICCS, volume 3994 of LNCS. Springer (2006). PP. 196–199.

[26] G. Sánchez, M. Rodriguez, "Cattle Marks Recognition by Hu and Legendre Invariant Moments". ARPN Journal of Engineering and Applied Sciences, Vol. 11, Nº 1, 2016.

[27] Li S., Lee M., and Pun C., "Complex Zernike Moments Features for Shape-Based Image Retrieval," IEEE Transaction on Systems, Man, and Cybernetics-Part A: Systems and Humans, vol. 39, no. 1, pp. 227-237, 2009.

Minería de Datos, Minería de Textos y Big Data

L. Lanzarini[1], W. Hasperué[1], A. Villa Monte[1,3], P. Jimbo Santana[4], G. Reyes Zambrano[5], J. Corvi[2], A. Fernandez Bariviera[6], J. A. Olivas[7]

[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] Facultad de Ciencias Administrativas, Universidad Central del Ecuador, Quito, Ecuador
[5] Facultad de Ciencias Físicas y Matemáticas, Universidad de Guayaquil, Guayaquil, Ecuador
[6] Dpto de Economía, Universitat Rovira i Virgili, Reus, España
[7] 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}@lidi.info.unlp.edu.ar
prjimbo@uce.edu.ec, gary.reyesz@ug.edu.ec, julieta.corvi@gmail.com , aurelio.fernandez@urv.net, cristina.puente@icai.comillas.edu, joseangel.olivas@uclm.es

CONTEXTO

Esta presentación corresponde a las tareas de investigación que se llevan a cabo en el III LIDI en el marco del proyecto "Sistemas inteligentes. Aplicaciones en reconocimiento de patrones, minería de datos y big data" perteneciente al Programa de Incentivos (2018-2021) y del proyecto PITAP-BA "Computación de Alto Desempeño, Minería de Datos y Aplicaciones de Interés Social en la Provincia de Bs.As." evaluado y subsidiado por la Comisión de Investigaciones Científicas de la Provincia de Bs.As. (2017-2019).

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 grandes volúmenes de textos y al procesamiento de flujo de datos.

En el área de la Minería de Datos se está trabajando, por un lado, en la construcción de conjuntos de reglas de clasificación difusas que faciliten y permitan justificar la toma de decisiones y, por otro lado, en el análisis de trayectorias vehiculares para predecir congestión de tránsito.

Con respecto al área de Big Data se está trabajando en el diseño y desarrollo de una técnica de clustering dinámico que se ejecuta de manera distribuida. Esta implementación se está llevando a cabo utilizando el framework Spark Streaming.

Por otro lado y como transferencia tecnológica concreta, se efectuó un análisis sobre la producción de leche en ganado bovino a partir de la base de datos de ARPECOL.

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, Stream processing.

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

Extracción de Reglas de Clasificación

Esta línea de trabajo comenzó hace varios años con el diseño de nuevos algoritmos para la obtención de conjuntos de reglas de clasificación haciendo énfasis en la simplicidad del modelo y la facilidad de interpretación por parte de quien debe tomar decisiones. En esa línea se definió oportunamente un algoritmo que demostró tener la capacidad de generar conjuntos de reglas adecuados [1]. Luego, con el objetivo de llevar a cabo una transferencia tecnológica en el área de Riesgo Crediticio se estudiaron consideraciones especiales basadas en información de expertos en otorgamiento de créditos. Se llegó a la conclusión de que era factible incorporar conjuntos difusos para tratar los atributos numéricos como variables lingüísticas y así simplificar la interpretación de las reglas por parte del oficial de crédito [2]. El nuevo método desarrollado se denomina FRvarPSO y tiene la capacidad de obtener de reglas de clasificación difusas operando sobre atributos nominales y numéricos [3]. Utiliza una técnica de optimización de población variable inicializada por medio de una red neuronal competitiva. Luego, por medio de los conjuntos difusos asociados a las variables lingüísticas, incorpora un criterio de votación que guía la búsqueda de las partículas facilitando la identificación de los valores a utilizar en la construcción las condiciones que darán lugar a los antecedentes de las reglas. Esta nueva propuesta fue medida en tres bases de datos de entidades financieras del Ecuador: una Cooperativa de Ahorro y Crédito y dos Bancos que otorgan diferentes tipos de crédito, con resultados satisfactorios [4]. Las mediciones realizadas utilizando 12 bases de datos del repositorio UCI y su comparación con otros métodos existentes en la literatura también han arrojados buenos resultados evidenciando la factibilidad de aplicar este nuevo método en distintos contextos [5].

Actualmente se sigue trabajando en el área de riesgo crediticio buscando mejorar dos aspectos: la incorporación de información macroeconómica y la incorporación de un factor de certeza a la recomendación dada por la regla. El primer aspecto debería impactar en la precisión de la regla y el segundo en la toma de decisiones por parte del agente de crédito.

Análisis de trayectorias GPS

El avance tecnológico facilita el registro y recolección de información de trayectorias GPS de vehículos en la vía pública. El análisis inteligente de estos datos lleva a identificar patrones sumamente útiles a la hora de tomar decisiones en situaciones relacionadas con urbanismo, circulación y congestionamiento, entre otras. En esta dirección se ha trabajado en el diseño e implementación de un nuevo método de agrupamiento de trayectorias GPS que utiliza información angular para segmentar los recorridos y una función de similitud guiada por un pivote. El proceso de adaptación inicia distribuyendo los centroides de manera uniforme en la región a analizar formando un reticulado. Los resultados obtenidos luego de aplicar el método propuesto sobre una base de datos de trayectorias reales fueron satisfactorios y muestran una mejoría significativa en comparación con los métodos publicados en la bibliografía.

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. Para el procesamiento en streaming se están desarrollando estrategias basadas en 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 [6].

Estas técnicas dinámicas se están implementando en el framework Spark Streaming, adecuado para procesamiento paralelo, distribuido y online.

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 [7].

Por otro lado y como una transferencia tecnológica concreta se está trabajando en el tratamiento de la información proveniente de ARPECOL, una asociación que nuclea entidades de control lechero de la provincia de Buenos Aires. Las entidades de control lechero son organizaciones que brindan el servicio de medición de la producción de leche individual a los productores. Estas entidades toman muestra de la leche de las vacas para realizar análisis de laboratorio de la calidad de la leche producida (porcentaje de grasa, de proteínas, de solidos totales, conteo de células somáticas). En esta línea de investigación se está colaborando con un proyecto del INIRA de la Facultad de Veterinaria de la UNLP que tiene por objetivo determinar factores genéticos para la identificación de las principales enfermedades reproductivas, mastitis y cojeras que afectan la lactancia de las vacas de tambo.

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 [8] 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 y (3) la incorporación de semánticas en el análisis [9]. 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 [10] 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 [11].

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.
- Modelización de trayectorias espacio-temporales con capacidad para establecer características comunes y detectar situaciones anómalas.
- Métodos estructurados y no estructurados aplicables a la representación de documentos.
- Representación de documentos de texto utilizando métricas.
- Obtención de resúmenes automáticos de texto.
- Implementación de técnicas inteligentes en el framework Spark Streaming
- Implementación de un algoritmo de clustering dinámico en Spark streaming.
- Análisis de la base de datos de ARPECOL para la identificación de características genéticas que mejoren la producción de leche de las vacas de tambo.

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.
- Diseño e implementación de un nuevo método de agrupamiento de trayectorias GPS aplicable a la predicción de congestiones vehiculares.
- 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 framework Spark streaming.
- 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.
- 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 doctores con dedicación exclusiva, 3 tesistas de Doctorado en Cs. Informáticas (1 becario doctoral UNLP), 1 tesista de grado y 2 profesores extranjeros.

Dentro de los temas involucrados en esta línea de investigación, en los últimos 2 años se han finalizado 2 tesis de doctorado y 5 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., Fernandez Bariviera A., Jimbo Santana P. *Simplifying Credit Scoring Rules using LVQ+PSO.* Kybernetes. Emerald Group Publishing Limited. vol. 46. pp 8-16. ISSN 0368-492X. 2017.

[2] 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.

[3] Jimbo Santana P., Lanzarini L., Fernández-Bariviera A. *Fuzzy Credit Risk Scoring Rules using FRvarPSO*. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems (ISSN 0218-4885). Vol 26. Nro1. pp. 39-57. World Scientific. 2018

[4] Jimbo Santana P., Lanzarini L., Fernández-Bariviera A. *Extraction of knowledge with population-based metaheuristics fuzzy rules applied to credit risk.* Advances in Swarm and Computational Intelligence. Lecture Notes in Computer Science. pp 153-163. Vol 10942. Springer, Cham. 2018.

[5] Jimbo Santana P., Lanzarini L., Fernández-Bariviera A. *FRvarPSO a method for obtaining fuzzy classification rules using optimization techniques*. International Conference on Modeling and Simulation in Engineering, Economics and Management (MS'2018). Girona. España (en prensa).

[6] Molina, R, Hasperué, W. *D3CAS: un Algoritmo de Clustering para el Procesamiento de Flujos de Datos en Spark*. XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). Octubre 2018.

[7] 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

[8] 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.

[9] 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 2019.

[10] Puente C., Villa Monte A., Lanzarini L., Sobrino A. and Olivas Varela J. Á. *Evaluation of causal sentences in automated summarie*. Proceedings of the 2017 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), págs. 1-6, doi. 10.1109/FUZZ-IEEE.2017.8015666, 2017.

[11] Coria, J.M. *Clasificación de Subjetividad utilizando Técnicas de Aprendizaje Automático*. Tesis de grado. Facultad de Informática, UNLP. Febrero 2018.

Minería de grafos en el dominio de trayectos académicos

Smail Ana[1], Pompei Sabrina[1], Bendati Natalia[2], Russo Claudia[3], Ramón Hugo[3], Fanny Vizcaino[2], Emanuel Lazzari[4]

Instituto de Investigación y Transferencia de Tecnología (ITT)[5]
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento Nro 1119 3er Piso, Junín (B) – TE: (0236) 4477050 INT 11610

{ana.smail, sabrina.pompei, natalia.bendati, claudia.russo, hugo.ramon, emanuel.lazzari}@itt.unnoba.edu.ar; fannybv8@gmail.com

RESUMEN

Entre las problemáticas actuales en la educación superior argentina se encuentra la necesidad de disminuir la duración real de las carreras. Hoy, los estudiantes universitarios "en promedio, tardan media carrera más de lo estipulado", afirma Mónica Marquina, coordinadora general del Programa de Calidad Universitaria de la Secretaría de Políticas Universitarias. [1]

Distintos factores provocan esta sobre-duración ya sea por causas vinculadas a condiciones extrínsecas a los estudiantes como intrínsecas a los mismos: desigualdad social, débil formación en la escuela secundaria, rigidez de los currículos, entre otros.

El presente trabajo propone inicialmente generar una base de datos orientada a grafos con la trayectoria académica de los egresados de las carreras Licenciatura en Sistemas e Ingeniería Informática de la UNNOBA. A partir del análisis de los grafos resultantes de dichas trayectorias, mediante el cálculo del camino más transitado y su comparación con el óptimo, se espera detectar posibles estructuras que incidan en la sobre-duración de las carreras en estudio.

En cuanto al diseño operativo del proyecto se propone una triangulación metodológica a fin de validar la información obtenida contra datos de entrevistas y encuestas a los propios egresados.

Esperamos que este trabajo nos permita avanzar en la línea de investigación minería de grafos, tanto en su aplicación como también en aspectos teóricos.

Palabras clave: minería de grafos, base de datos orientada a grafos, trayectoria académica, duración real de carreras

---

[1] Docente Investigador - ITT
[2] Becario - ITT
[3] Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC
[4] Becario CIN - ITT
[5] ITT - Centro Asociado CIC

CONTEXTO

La línea descripta está inserta en los proyectos de investigación “Informática y Tecnologías Emergentes" y “Tecnología y aplicaciones de Sistemas de Software: Innovación en procesos, productos y servicios”, presentados en la convocatoria “Subsidios de investigación bianuales” 2019 de la Universidad Nacional del Noroeste de Buenos Aires.

Dicho proyecto tiene lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) de la UNNOBA, Centro Asociado CIC. Dada la dependencia académica del mismo con la Escuela de Tecnología (ET) de la UNNOBA, se propone el estudio de duración de las carreras del área informática correspondiente a la oferta académica de la ET.

1.INTRODUCCIÓN

La creciente demanda de análisis basado en relaciones ha puesto en escena el uso de grafos para el descubrimiento de conocimiento. El modelo en grafo es útil particularmente cuando los datos a almacenar tienen multitud de interrelaciones entre sí, y cuando la importancia recae más en las interrelaciones que se establecen entre los datos, que en los propios datos. La minería de grafos es la encargada de encontrar patrones significativos, útiles y novedosos en una representación basada en grafos de los datos originales.

Un grafo G está compuesto por un conjunto de vértices V(G) y un conjunto de enlaces o arcos E(G). Un enlace es un par no ordenado de vértices que se denota {u, v}. Un grafo G’ = (V’, E’) es subgrafo de G si V’ □ V y E′□ E y se denota G’ □ G .

Instancia		Grafo
Elemento	⇔	vértice
Atributos Elemento	⇔	Etiqueta Vértice
Relaciones	⇔	Arcos
Tipo de Relaciones	⇔	Etiqueta Arcos

Una representación basada en grafos es lo suficientemente flexible para permitir tener más de una representación para un dominio dado. Esto permite al investigador experimentar para obtener la mejor representación para su dominio. [2] Las bases de datos pueden ser conjuntos de grafos de pequeño o mediano tamaño o un único grafo conexo de gran tamaño. En base a esto, las consultas más realizadas se consisten en buscar subgrafos similares a una estructura dada o, a partir de una frecuencia de ocurrencia determinada, buscar todos los subgrafos que se repitan con igual o mayor frecuencia entre todos los grafos de la base o dentro de un grafo de gran tamaño. [3]

La minería de datos basada en grafos no solo se enfoca en encontrar subestructuras repetitivas y comunes dentro de los datos de entrada, sino también en el proceso de identificar conceptos que describen a las subestructuras más importantes para una mejor interpretación de los datos. Una vez descubierta la misma podrá ser utilizada para simplificar el grafo original mediante el reemplazo de la subestructura por un vértice que represente a la recién descubierta subestructura.

En cuanto al espacio de búsqueda de un algoritmo basado en grafos, este está compuesto de todos los grafos derivados a partir del grafo de entrada. Es decir, la búsqueda de patrones discriminativos necesita realizar pruebas de isomorfismo de subgrafos en los grafos de entrada, el cual es un problema NP-Completo. [4]

A fin de reducir la complejidad algunos algoritmos dividen los datos de entrada en varios grupos y de esta manera el número de comparaciones entre los elementos se reduce. De hecho, existen varios algoritmos de minería de datos, entre los cuales se encuentran el FSG, FFSM, gSpan y GASTON. [5] Los mismos

varían en la estrategia que utilizan para recorrer los grafos, el tipo de entrada que utilizan y la información de salida que proveen. Para dar soluciones más eficientes a los problemas necesario compararlos y determinar cuál es el más apropiado para el dominio de datos con los que se cuenta.

Sobre el espacio de búsqueda es posible además observar distintos parámetros y/o métricas que proporcionan información sobre el dominio en estudio. Por ejemplo: la centralidad de grado, centralidad de cercanía, centralidad de intemediación; componentes fuertemente conectados, componentes débilmente conectados, tamaño de componente gigante; rutas más cortas; densidad del grafo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación propuestas se enmarcan en la minería de grafos.

- Exploración de posibles representaciones para las trayectorias académicas mediante grafos.
- Generación de una base de datos orientada a grafos para almacenar trayectorias académicas.
- Estudio y evaluación de algoritmos para la detección de patrones en el dominio de los trayectos académicos.
- Estudio de herramientas de visualización y análisis de grafos.

3.RESULTADOS OBTENIDOS/ESPERADOS

Mediante la aplicación de minería de grafo al dominio de los trayectos académicos de los egresados de las carreras informáticas de la UNNOBA, se espera detectar posibles factores que inciden en la duración real de las carreras. Esta información puede ser clave para la toma de decisiones a fin de disminuir el índice de retraso en la carrera.

Se espera, además, alcanzar un modelo de grafo conceptual aplicable a otras carreras de la Universidad.

Para ello proponemos los siguientes objetivos específicos:

- Explorar patrones topológicos de grafos que contengan información relevante sobre los trayectos académicos de los egresados de las carreras informáticas de la UNNOBA.
- Analizar la existencia de patrones.
- Aplicar algoritmos que permitan obtener el camino más transitado.
- Comparar el camino más transitado con el camino óptimo.
- Inferir posibles causas de retraso en la carrera.
- Comparar la información obtenida con los datos obtenidos de entrevistas y encuestas de los egresados.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo está compuesto por docentes investigadores, becarios alumnos, y alumnos avanzados de las carreras del área de informática.

En el marco del proyecto se desarrollarán dos tesis de maestría, una en minería de datos y otra en metodologías de la investigación científica. Como así también una tesina de grado correspondiente a la Licenciatura en Sistemas.

5. BIBLIOGRAFÍA

[1] “Reconocimiento académico: pensar la formación con foco en el estudiante.”, Ministerio de Educación, Cultura, Ciencia y Tecnología, 2018. [En línea]. Disponible en: https://www.argentina.gob.ar/noticias/reconocimiento-academico-pensar-la-formacion-con-foco-en-el-estudiante [Accedido: 12-mar-2019]

[2] A. Cortez Vásquez y L. Pro Concepción, “Descubrimiento de conocimiento basado en

grafos", *Revista de Investigación de Sistemas e Informática*, vol 8, no. 2, pp 17-25, 2011 [En línea]. Disponible en: http://revistasinvestigacion.unmsm.edu.pe/index.php/sistem/article/download/6321/5541 [Accedido: 26-feb-2019]

[3] S. Bianco, "Análisis comparativo de algoritmos de minería de subgrafos frecuentes", tesis de grado, Universidad Nacional de Lanus, 2016 [En línea]. Disponible en: http://sistemas.unla.edu.ar/sistemas/gisi/TFLS/Bianco-TFL.pdf. [Accedido: 1-mar-2019]

[4] R. Salomón Fonseca Delgado, "Diseño de un algoritmo de minería de datos basado en grafos para la tarea de aprendizaje de conceptos", tesis de maestría, Instituto Nacional de Astrofísica, Optica y Electrónica de Puebla, 2012 [En línea]. Disponible en: https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/745/1/FonsecaDR.pdf [Accedido: 26-feb-2019]

[5] S. Bianco, S. Martins y R. García Martinez, "Estudio compartivo de algoritmos de minería de subgrafos frecuentes" en 2016 XXII Congreso Argentino de Ciencias de la Computación, pp 748-457 [En línea]. Disponible en: http://hdl.handle.net/10915/56756 [Accedido: 3-mar-2019]

- P. Tan, M. Steinbach y V. Kumar, *Introduction to Data Mining*. New York, NY, USA: Pearson, 2005.
- T. Munzner, *Visualization Analysis & Design*. Boca Raton, USA: CRC Press, 2014.
- I. Robinson, J. Webber, and E. Eifrem, *Graph Databases*. Sebastopol, USO: O'Reilly, 2015.
- L. Haberfeld, M. Marquina y S. Morresi, *El Sistema universitario argentino. Situación, problemas y políticas*. Centro de Estudios para el cambio estructural (CECE), 2018. [En línea] Disponible en: http://fcece.org.ar/el-sistema-universitario-argentino/ [Accedido: 14-nov-2018]

Modelos para Aprendizaje Automático en Tiempo Real sobre Entornos de Big Data

Banchero Santiago, Fernandez Juan M., Tonin Monzón Francisco
Giordano Luis A., Marrone Agustín H., Paz Soldan Carlos
Tolosa Gabriel H.
{sbanchero, jmfernandez, ftonin, agiordano, amarrone, cpazsoldan, tolosoft}@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 que permitan 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 han acrecentado el interés por la utilización de técnicas tradicionales de aprendizaje automático en distintos problemas de repositorios masivos y entornos de flujos (o *streaming*) de datos donde muchas veces no es posible su almacenamiento, pero se requiere tomar decisiones en tiempo real.

Contexto

Este proyecto es el comienzo de una nueva línea de investigación del Departamento de Ciencias Básicas (UNLu) que tiene como principal aspiración 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) y financieras (por ejemplo: banca y megatiendas entre otros) son algunas de sus aplicaciones [1, 13]. Esta problemática se conoce como el problema de Big Data [16].

Big Data se caracteriza principalmente 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 [12]. Si bien el volumen es y será un desafío significativo del Big Data, se debe prestar mucha atención a todas las dimensiones del problema: Volumen, Variedad y Velocidad (conocidos como las 3Vs) [6].

El concepto de *streaming* en Big Data tiene algunas características distintivas, 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: tráfi-

co de red, redes de sensores, cuidado de la salud, etc.), seguimientos de *clics* 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 [14]. Los entornos de *streaming* de datos proponen nuevos retos con respecto a estas etapas. Una característica importante en selección de *features* es la habilidad para manejar grandes volúmenes de datos [18, 20]. 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. [11, 23, 22].

Además del problema de selección de *features*, hay cuestiones enmarcadas dentro de las tareas de *stream 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, 21]. Incluso la distribución subyacente puede cambiar en el tiempo, lo que genera que algunos modelos ya no sigan siendo válidos. Estos aspectos hacen 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 las diferentes aplicaciones a datos reales.

Líneas de I+D

En este proyecto se inician lineas 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. Asimismo, se están analizando diferentes opciones de topologías con Storm Apache[1] para la resolución de problemas de ETL (*Extract, Transform and Load*). A su vez, en el último tiempo se ha incorporado una nueva linea de investigación que refiere al estudio de algoritmos de clasificación multi-etiquetas, también en ambientes de *streaming* de datos. A continuación se hace una descripción somera de estas líneas de I+D.

a. Árboles de decisión adaptativos

Los árboles de decisión corresponden al aprendizaje supervisado y son ampliamente utilizados en problemas de clasificación. Estos algoritmos intentan, a partir de las instancias vistas, generar hipotesis con las cuales hacer predicciones de futuras instancias [15].

En aprendizaje sobre flujo de datos (*stream learning*), en general no es necesario computar estadísticas sobre todo el pasado, siendo suficiente con hacerlo sobre el pasado reciente [9]. Una de las formas mas clásicas y simples de mantener los ejemplos correspondientes a ese pasado es almacenar solo una ventana de instancias.

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].

[1] `http://storm.apache.org/`

Los resultados preliminares de aplicar HAT en el dominio de gestión de caché de consultas en motores de búsqueda han sido muy alentadores [19]. Como objetivo principal, se propuso evaluar la performance de un árbol de decisión adaptativo (HAT) para aplicar al diseño de una política de admisión para un motor de búsqueda web que recibe (y procesa) consultas en modo *streaming*. Se ha trabajado modelando la admisión como un problema de clasicación binario, intentando capturar los cambios de concepto en el tiempo, manteniendo un modelo siempre ajustado. Una vez ajustado el modelo, se integró el mecanismo de decisión como política de admisión y se evaluó la performance del caché de resultados, comparando el modelo resultante de utilizar árboles de decisión adaptativos con algoritmos de clasificación tradicionales. Siendo, de acuerdo a nuestro conocimiento, la primera vez que se propone un árbol de decisión adaptativo para la detección de términos de búsqueda frecuentes en motores de búsqueda, en los experimentos realizados, se ha observado un incremento del rendimiento del 18 % en comparación con la utilización de técnicas de clasificación tradicionales.

El 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 conjuntos 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 descubrimiento de conocimiento, ha representado un gran desafío durante los últimos años tanto en estadística como en problemas de aprendizaje automático [22]. 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* en 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[2] 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.

d. Clasificación multi-etiquetas

La clasificación multi-etiquetas es un nuevo paradigma de aprendizaje supervisado que generaliza

[2] `http://storm.apache.org/`

las técnicas clásicas de clasificación para abordar problemas en donde cada instancia de una colección se encuentra asociada a múltiples etiquetas [10].

La mayor parte de los trabajos de investigación en este campo han sido realizados en contextos de aprendizaje por *batch* [8]; sin embargo, los ambientes de flujo continuo de datos (o streaming) presentan nuevos desafíos debido a las limitaciones de tiempo de respuesta y almacenamiento que acarrean. A esto se agrega la naturaleza evolutiva de este tipo de escenarios, que obligan a los algoritmos a adaptarse a cambios de concepto [17].

Una propuesta para esta linea de investigación sugiere aplicar algoritmos de clasificación multietiquetas a colecciones estructuradas y no estructuradas, combinando estos algoritmos con técnicas de procesamiento de lenguaje natural sobre la colección no estructurada. A su vez, por último, se proponen abordar estrategias de ensambles de algoritmos en búsqueda de una mejora en la calidad de la tarea de predicción de objetos no observados por el modelo.

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. Elaborar metodologías para el desarrollo de modelos en linea 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. Recientemente se ha finalizado un trabajo final correspondiente a la Lic. en Sistemas de Información (UNLu), se están dirigiendo dos más y se espera dirigir al menos dos por año hasta la finalización del proyecto. Por otro lado, se está dirigiendo una beca de investigación EVC y se espera presentar dos candidatos más 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] Gama, J., and Gaber, M. M. *Learning from data streams: processing techniques in sensor networks*. Springer, 2007.

[10] Gibaja, E., and Ventura, S. A tutorial on multilabel learning. *ACM Computing Surveys (CSUR) 47*, 3 (2015), 52.

[11] Huang, H., Yoo, S., and Prasad, S. Unsupervised Feature Selection on Data Streams. 1031–1040.

[12] Khan, N., Yaqoob, 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.

[13] Marz, N., and Warren, J. *Big Data: Principles and best practices of scalable realtime data systems*. Manning Publications Co., 2015.

[14] Pruengkarn, 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.

[15] Quinlan, J. R. Induction of decision trees. *Machine learning 1*, 1 (1986), 81–106.

[16] Safaei, A. A. Real-time processing of streaming big data. *Real-Time Systems 53*, 1 (2017), 1–44. 00004.

[17] Sousa, R., and Gama, J. Multi-label classification from high-speed data streams with adaptive model rules and random rules. *Progress in Artificial Intelligence 7*, 3 (2018), 177–187.

[18] Tang, J., Alelyani, S., and Liu, H. Feature selection for classification: A review. *Data Classification: Algorithms and Applications* (2014), 37.

[19] Tonin Monzón, F., Banchero, S., and Tolosa, G. H. Árboles de decisión adaptativos en políticas de admisión a caché. In *IV Simposio Argentino de GRANdes DAtos (AGRANDA 2018)-JAIIO 47 (CABA, 2018)* (2018).

[20] 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.

[21] Wang, L. Machine learning in big data. *International Journal of Advances in Applied Sciences 4*, 4 (2016), 117–123. 00048.

[22] 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.

[23] Zhichao, Y., and Chunyong Yin, L. F. A Feature Selection Algorithm of Dynamic Data-Stream Based on Hoeffding Inequality. 92–95.

Modelos y Algoritmos para Problemas de Procesamiento en Entornos de Big Data

Gabriel H. Tolosa[1], Pablo Lavallén[1], Tomás Delvechio[1], Agustín Marrone[1], Andrés Giordano[1] , Esteban A. Ríssola[1,3]
{tolosoft, plavallen, tdelvechio, eamarrone, agiordano}@unlu.edu.ar
esteban.andres.rissola@usi.ch

[1]Departamento de Ciencias Básicas, Universidad Nacional de Luján
[2]Facoltà di Scienze Informatiche, Università della Svizzera Italiana

Resumen

La idea del procesamiento de datos masivos (Big Data) se ha desarrollado sostenidamente en los últimos años, estableciéndose como un nuevo paradigma para resolver problemas. Por un lado, el crecimiento en el poder de cómputo y almacenamiento habilita la posibilidad de manejar volúmenes de datos de varios órdenes de magnitud. Por el otro, generan la necesidad de contar no solamente con plataformas que permitan distribuir el procesamiento sino, además, con algoritmos que lo realicen de forma eficiente.

Una de las primeras aplicaciones de Big Data son los motores de búsqueda de escala web, sistemas que procesan miles de millones de documentos y deben responder a los usuarios con estrictas restricciones de tiempo, típicamente, milisegundos. Análogamente, el procesamiento de grafos masivos provenientes del mapeo de la estructura de las redes sociales presenta desafíos de forma sostenida.

Estos escenarios se caracterizan por una complejidad creciente en espacio y requieren soluciones cada vez más sofisticadas ya que la cantidad de datos y de usuarios crece conforme evolucionan en el tiempo. Además, han aparecido requerimientos para ofrecer respuestas sobre flujos de datos que ocurren en tiempo real (*streaming*) por lo que es un requisito considerar modelos que puedan tomar decisiones *on-line* utilizando estos datos.

Este trabajo presenta las líneas de investigación que se proponen en el contexto de los datos masivos a partir del estudio, diseño y evaluación de estructuras de datos y algoritmos que operan eficientemente, ya sea sobre documentos, grafos sociales o interacciones de usuarios, entre otros.

Palabras clave: algoritmos eficientes, motores de búsqueda, grafos masivos, estructuras de datos, grandes datos.

Contexto

Esta presentación se encuentra enmarcada en el proyecto de investigación "Algoritmos Eficientes y Minería Web para Recuperación de Información a Gran Escala" del Departamento de Ciencias Básicas (UNLu).

Introducción

La gran asimilación de tecnologías de la información y la comunicación por parte de las sociedades en las últimas décadas genera un renovado interés sobre algoritmos que permitan almacenar y procesar grandes volúmenes de datos en límites de tiempo cada vez más precisos [28]. Además se observa un auge y maduración de tecnologías de cómputo en la nube, base sobre la cual se plantea la posibilidad de acceder a crecientes capacidades de procesamiento y almacenamiento de forma rápida y con esfuerzo reducido en el despliegue y gestión de recursos. Una de las áreas donde se dio la convergencia de ambos procesos generando nuevas oportunidades de investigación es el área de Big Data (Datos Masivos). Este enfoque ofrece esquemas de procesamiento y almacenamiento distribuido considerando prioritario permitir la escalabilidad horizontal y la tolerancia a fallas, como es el caso de la plataforma Spark [26].

De forma complementaria, surge la necesidad permanente de nuevos modelos y estrategias para resolver los problemas que aparecen continuamente

de forma eficiente. Un caso clásico son los motores de búsqueda web, que procesan miles de millones de documentos[1] y deben responder a los usuarios con estrictas restricciones de tiempo, típicamente, milisegundos [2]. Otro ejemplo común hoy en día es el procesamiento de grafos masivos provenientes, por ejemplo, de la estructura de una red social [18] o de interacciones entre usuarios. También son ejemplos los servicios de *streaming* [28, 22], la genómica [20] y la meteorología [15].

Estos problemas requieren, en general, procesamiento distribuido, paralelo y algoritmos altamente eficientes [5]. Para su abordaje, se utilizan plataformas de almacenamiento y procesamiento no tradicionales [17], como Hadoop [25] y Spark [26], sobre las cuales se situan capas de software que implementan algoritmos de búsquedas, aprendizaje automático y optimización. En la mayoría de los casos, la partición del problema y la distribución de la carga de trabajo son aspectos de las estrategias que requieren ser optimizados de acuerdo al problema [24].

Líneas de I+D

En este trabajo se describen líneas de I+D del grupo, las cuales se basan, principalmente, en mejorar la eficiencia en la recuperación de información de gran escala y el procesamiento de grandes volúmenes de datos, en general, sobre *commodity hardware* [9]. De forma continua aparecen nuevos desafíos en estas áreas y 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

a.1. Algoritmos para *Top-k*

En la actualidad, lo común es que los motores de búsqueda respondan en cuestión de milisegundos a las consultas de los usuarios (*query*), con un conjunto de los k documentos más relevantes. Para ello, deben revisar una colección que se compone de miles de millones de elementos, por lo que es crucial mejorar y explorar nuevas técnicas que permitan "atravesar" de forma eficiente las estructuras de datos internas. A los algoritmos que toman un *query* y revisan un índice invertido en búsqueda de los documentos más relevantes se los pueden agrupar en dos categorías: i) TAAT (term-at-a-time), procesan de a un término del query a la vez, y ii) DAAT (document-at-a-time), en los que se procesa de a un documento a la vez. Para ello, en la familia de algoritmos DAAT, MaxScore [23] y WAND [13] almacenan un *score* global máximo (*upperbound*) y lo utilizan para saber si un documento puede formar parte del resultado final en los *top-k*. Como extensión a estos, existen los algoritmos Block-Max WAND [10], Variable Block-Max WAND [19] y Block-Max MaxScore [6], los cuales almacenan no sólo uno, sino un *upperbound* local a un bloque de cierto tamaño, con el fin de procesar aún una cantidad de entradas aún menor.

En esta línea se investigan los algoritmos DAAT que forman parte del estado del arte y, adicionalmente, se propone uno nuevo como extensión de MaxScore. En éste, en vez de almacenar sólo un *upperbound* global, se almacena un conjunto de ellos en una estructura similar a una *skip list* [7], dotando al algoritmo de más información para mejorar la eficiencia del procesamiento.

a.2. Búsquedas Selectivas

Complementando los algoritmos de búsqueda mencionados en la línea anterior, otra estrategia para acelerar el proceso está basada en la partición de la colección de documentos en porciones, P, (*shards*) de acuerdo a algún criterio (por ejemplo, temático) de manera tal de enviar las consultas solamente a un número reducido n de nodos ($n << P$) que contengan particiones de la colección que potencialmente pueden satisfacer la consulta. Este problema se lo conoce como "búsquedas selectivas"(*selective search*) [16] e incluye métodos que permiten seleccionar los recursos adecuados, algoritmos de fusión de resultados parciales y estrategias adaptadas de *caching*.

Dichas estrategias son desafiadas cuando los documentos aparecen en flujos en tiempo real como, por ejemplo, las publicaciones en las redes sociales. Un caso paradigmático son las publicaciones en Twitter, en la cual millones de usuarios alrededor del mundo publican "documentos cortos"(*tweets*) desde diferentes tipos de dispositivos (generalmente, móviles), los cuales deben estar disponibles casi de inmediato (segundos) por lo que las estructuras de datos deben soportar un alto dinamismo [4].

[1] Según http://www.worldwidewebsize.com el índice de Google contiene aproximadamente unos 55.000 millones de documentos (Marzo/2019).

En esta linea de investigación se propone combinar la problemática de las búsquedas en tiempo real utilizando una arquitectura basada en búsquedas selectivas donde los criterios de actualización del índice invertidos por partición, las estrategias de caché a implementar y el algoritmo de búsqueda final impactan en la performance que se pretende optimizar (eficiencia y/o efectividad). Los objetivos generales perseguidos son el desarrollo de técnicas y algoritmos para la asignación de flujos de documentos a diferentes particiones de un índice para luego soportar búsquedas sobre un subconjunto de éstas, maximizando la performance (tanto eficiencia como efectividad).

b. Procesamiento de Grafos Masivos

b.1. Cálculo Distribuido en Grafos Evolutivos

Como se ha dicho anteriormente, uno de los desafíos surgidos con la aparición de las aplicaciones que procesan datos masivos es la de generar representaciones adecuadas para los datos recolectados. Una de las estructuras más longevas y versátiles de las ciencias de la computación para estas representaciones son los grafos, formados por un conjunto de nodos y relaciones entre éstos (aristas) [7]. Diversos algoritmos que se aplican sobre grafos son ampliamente utilizados en la industria y la academia para abordar un gran abanico de soluciones. Algunos de estos algoritmos son los de camino mas corto (Single Source Shortest Path o SSSP y All-Pairs Shortest Path o APSP), y suelen constituir la base de muchos procesos en ámbitos como la recuperación de información [3] o análisis de redes sociales [1], por nombrar un par de ejemplos.

Este tipo de procesos suelen tener soluciones aceptadas desde hace mucho tiempo, las cuales no escalan cuando el tamaño de los datos supera un umbral, dando lugar a soluciones en tiempos no triviales [8]. Por otro lado, es todo un desafío la implementación de estrategias que aprovechen las características distribuidas de las plataformas existentes [12]. Otro problema que presentan los enfoques secuenciales tradicionales y no escalables es que muchas aplicaciones exigen de forma creciente respuestas cada vez mas instantáneas, lo que introduce la posibilidad de construir soluciones aproximadas en el resultado pero eficientes en el tiempo insumido. Otro punto a considerar es que los grafos, al representar relaciones, pueden cambiar a lo largo del tiempo, que se conoce como grafos dinámicos [14], y admiten soluciones que tengan en cuenta procesamientos previos para no tener que recalcular completamente las métricas ante cada modificación del mismo.

Dentro del marco del proyecto, el objetivo de esta linea de trabajo es diseñar, analizar e implementar algoritmos eficientes para el cálculo de las distancias mínimas (*shortest path*) para grafos que pueden o no evolucionar en el tiempo. Estas soluciones se abordan utilizando estrategias de cómputo distribuido sobre plataformas de Big Data (por ejemplo, Spark) sobre hardware commodity y soportando tolerancia a fallas.

b.2. Estimación de Distancia entre Nodos

El problema de la distancia entre dos nodos tiene múltiple aplicaciones prácticas, por ejemplo, para el ranking en búsquedas[2]. Se lo define como la longitud del camino más corto entre ellos y se vuelve 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 [27] y agrupamiento de usuarios [11], entre otros.

El cálculo exacto de esta métrica entre dos nodos arbitrarios se ve afectado en cuanto a la eficiencia del proceso debido al tamaño de los grafos actuales resultando prohibitivo para aplicaciones prácticas. Por ejemplo, en redes sociales digitales el número de relaciones (aristas) supera ampliamente la cantidad de usuarios (nodos)[3]

La estimación de este valor es una alternativa válida y la reducción del error asociado (a un costo computacional bajo) conforma una de las líneas de investigación de este proyecto. En este enfoque se utiliza un conjunto de nodos, llamados *landmarks* [21], 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. Algunos resultados preliminares sobre nuevos métodos de selección

[2] Un caso es la red de contactos profesionales Linkedin.

[3] Twitter: 321 millones de usuarios activos, 226.947 millones de vínculos. `http://investor.twitterinc.com/home/default.aspx`

de *landmarks* indican que usar métricas tradicionales en conjunto (*Closeness Centrality*, grado, entre otras) , sobre ciertos grafos, potencian la estimación de la distancia logrando mejores resultados respecto del uso individual de éstas.

El desarrollo de métodos de corrección de la distancia estimada no ha sido abordado hasta el momento y es otra de las lineas de investigación de este proyecto. La idea consiste en lograr un algoritmo que, basado en características generales del grafo (como la densidad, diámetro, tamaño de comunidades, entre otras) permita corregir el valor de la estimación. Algunos análisis preliminares sobre grafos sintéticos indican que es posible reducir el error mediante esta técnica.

Resultados y objetivos

El objetivo principal del proyecto es desarrollar, evaluar y transferir modelos y algoritmos para abordar algunas de las problemáticas relacionadas con las búsquedas a gran escala y el procesamiento de grandes datos, en este caso, grafos masivos. En general, se proponen mejoras que apuntan a la eficiencia y la escalabilidad. En particular, se espera alcanzar los siguientes objetivos:

- Diseñar versiones optimizadas de algoritmos de procesamiento de *queries* dotando a los algoritmos para $top-k$ de información accesoria que les permita recorrer las estructuras de datos eficientemente.
- Desarrollar de técnicas y algoritmos para la asignación de flujos de documentos a diferentes particiones de un índice para luego soportar búsquedas selectivas sobre un subconjunto de éstas, maximizando la performance (tanto eficiencia como efectividad) y considerando parámetros de la arquitectura (por ejemplo, número de procesadores, núcleos, etc.).
- Definir y evaluar estructuras y modelos de cómputo distribuido sobre hardware commodity para problemas de cálculo de métricas en grafos masivos, en particular en grafos evolutivos.
- Diseñar y evaluar estrategias de estimación de distancias entre nodos de un grafo masivo para problemas de búsqueda, junto con métodos de corrección de la estimación.

Formación de Recursos Humanos

En el marco de estas líneas de investigación se están dirigiendo dos tesis de Licenciatura en Sistemas de Información (UNLu). Además, asociados al proyecto de investigación hay una estancia de investigación de la Secretaría de CyT (UNLu), una Beca Estímulo a las Vocaciones Científicas (CIN) y dos pasantías internas UNLu.

Referencias

[1] D. A. Bader, S. Kintali, K. Madduri, and M. Mihail. Approximating betweenness centrality. In *International Workshop on Algorithms and Models for the Web-Graph*, pages 124–137. Springer, 2007.

[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] S. Brin and L. Page. The anatomy of a large-scale hypertextual web search engine. *Computer networks and ISDN systems*, 30(1-7):107–117, 1998.

[4] M. Busch, K. Gade, B. Larson, P. Lok, S. Luckenbill, 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] K. Chakrabarti, S. Chaudhuri, and V. Ganti. Interval-based pruning for top-k processing over compressed lists. In *Proc. of the 2011 IEEE 27th International Conference on Data Engineering*, ICDE '11, pages 709–720, USA, 2011. IEEE Computer Society.

[7] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to algorithms.* MIT press, 2009.

[8] A. Crauser, K. Mehlhorn, U. Meyer, and P. Sanders. A parallelization of dijkstra's shor-

test path algorithm. In *International Symposium on Mathematical Foundations of Computer Science*, pages 722–731. Springer, 1998.

[9] T. Delvechio and G. H. Tolosa. Indexación distribuida con restricción de recursos. In *Simposio Argentino de GRANdes DAtos - JAIIO 46 (Córdoba, 2017)*, 2017.

[10] 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.

[11] J. Edachery, A. Sen, and F. J. Brandenburg. Graph clustering using distance-k cliques. In J. Kratochvíyl, editor, *Graph Drawing*, Berlin, Heidelberg, 1999. Springer Berlin Heidelberg.

[12] N. Edmonds, T. Hoefler, and A. Lumsdaine. A space-efficient parallel algorithm for computing betweenness centrality in distributed memory. In *2010 International Conference on High Performance Computing*, pages 1–10. IEEE, 2010.

[13] A. B. et al. Efficient query evaluation using a two-level retrieval process. *Proc. Proc. of the twelfth international conference on Information and knowledge management, ACM*, pages 426–434, 2003.

[14] D. Ferone, P. Festa, A. Napoletano, and T. Pastore. Shortest paths on dynamic graphs: A survey. *Pesquisa Operacional*, 37(3):487–508, 2017.

[15] X. Guo. Application of meteorological big data. In *Communications and Information Technologies (ISCIT), 2016 16th International Symposium on*, pages 273–279. IEEE, 2016.

[16] A. Kulkarni and J. Callan. Selective search: Efficient and effective search of large textual collections. *ACM Transactions on Information Systems (TOIS)*, 33(4):17, 2015.

[17] S. Madden. From databases to big data. *IEEE Internet Computing*, 16(3):4–6, 2012.

[18] J. Magnusson. Social network analysis utilizing big data technology, 2012.

[19] A. Mallia, G. Ottaviano, E. Porciani, N. Tonellotto, and R. Venturini. Faster blockmax wand with variable-sized blocks. In *Proc. of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '17, pages 625–634, New York, NY, USA, 2017. ACM.

[20] 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.

[21] M. Potamias, F. Bonchi, C. Castillo, and A. Gionis. Fast shortest path distance estimation in large networks. In *Proc. of the 18th ACM Conference on Information and Knowledge Management*, CIKM '09, pages 867–876, New York, NY, USA, 2009. ACM.

[22] E. Ríssola 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.

[23] H. R. Turtle and J. Flood. Query evaluation: Strategies and optimizations. *Inf. Process. Manage.*, 31(6):831–850, 1995.

[24] Y. Wang, L. Wu, L. Luo, Y. Zhang, and G. Dong. Short-term internet search using makes people rely on search engines when facing unknown issues. *PloS one*, 12(4):e0176325, 2017.

[25] T. White. *Hadoop: The Definitive Guide*. O'Reilly Media, Inc., 1st edition, 2009.

[26] 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.

[27] 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.

[28] P. Zikopoulos, C. Eaton, et al. *Understanding big data: Analytics for enterprise class hadoop and streaming data*. McGraw-Hill Osborne Media, 2011.

Recuperación de Datos para el Procesamiento de Datos Masivos

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
{*fkasian, vlud, mprinti, nreyes,proggero*}@unsl.edu.ar, merenda.franco83@gmail.com
Karina Figueroa
Universidad Michoacana de San Nicolás de Hidalgo, México
karina@fismat.umich.mx
Claudia Deco
Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario
deco@fceia.unr.edu.ar

Resumen

La gran cantidad y variedad de información disponible en formato digital, junto con la evolución de las tecnologías de información y comunicación, han llevado al surgimiento de nuevos repositorios no estructurados de información, en los cuales los datos son no estructurados y 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 fácilmente en claves y registros (tanto manual como computacionalmente), o tal estructuración carece de sentido práctico , y restringe de antemano los tipos de consultas que luego se pueden realizar. Por lo tanto, es cada vez más evidente en la actualidad la necesidad de procesar grandes conjuntos de datos, para obtener información útil a partir de ellos.

Así, dada una base de datos y una consulta, el objetivo de un sistema de recuperación de información es obtener, desde la base de datos, lo que podría ser útil o relevante para el usuario, usando alguna estructura de almacenamiento sobre dichos datos que permita responder a la consulta de manera eficiente. Estas estructuras necesitan ser diseñadas especialmente para ese propósito.

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 3-03-2018 de la Universidad Nacional de San Luis (UNSL) y en el Programa de Incentivos (Código 22/F834): "Tecnologías Avanzadas Aplicadas al Procesamiento de Datos Masivos", dentro de la línea "Recuperación de Datos e Información", desarrollada en el Laboratorio de Investigación y Desarrollo en Inteligencia Computacional (LIDIC) de la UNSL. Este proyecto ha sido aprobado en la UNSL por Resolución del Consejo Superior 126/18 y finaliza en 2021.

En esta línea de investigación se busca desarrollar herramientas eficientes para sistemas de recuperación 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 este tipo de datos.

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. , favoreciendo en ellos la disponibilidad de estructuras de datos eficientes y escalables, para memorias jerárquicas, que además aprovechen, cuando sea necesario, la aplicación de técnicas de computación de alto desempeño (HPC).

2. Introducción y Motivación

En nuestros días, el uso masivo de internet y la disponibilidad de dispositivos electrónicos en diversos ámbitos, ha generado una significativa aceleración tanto en el crecimiento del volumen de datos capturados y almacenados, como la variedad de tipos de datos que aparecen. En este escenario, surge la necesidad de redefinir las técnicas tradicionales para el procesamiento, análisis y obtención de información útil, para formular nuevas metodologías de abordaje y lograr una mayor aplicabilidad.

Como se sabe, aún en la actualidad, los sistemas tradicionales de computación utilizan principalmente información estructurada. Este tipo de información puede organizarse en claves y registros, sobre los cuales las búsquedas tradicionales tienen sentido y donde la estructura misma de los datos puede interpretarse y utilizarse en programas casi directamente. Sin embargo, como se ha mencionado, dado el volumen y variedad de los datos actualmente disponibles, dos de las características que aparecen en los datos en el contexto de problemas de "big data", no es posible restringir las búsquedas sobre datos estructurados a las búsquedas tradicionales, 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 mucho más generales, que puedan servir de apoyo, por ejemplo, en la toma de decisiones. Las búsquedas por similitud son un tipo de búsqueda más general, que se sustentan, para lograr eficiencia en las respuestas, sobre *métodos de acceso* o *índices métricos* [5].

Por ello, si se consideran conjuntos masivos de datos no estructurados, dada la gran cantidad de datos con los que se trabaja, se pueden utilizar estos índices para lograr eficiencia en la respuesta cuando se presentan al sistema consultas de recuperación de información . Dichos índices pueden tener distintas características que los hacen adecuados 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". Dicho tipo de búsqueda usa el dato no estructurado en sí mismo para describir lo que se busca. La similitud entre dos objetos se calcula mediante una función de distancia, la cual intenta describir realmente la disimilitud entre ellos.

El modelo habitual para las búsquedas por similitud es el de espacios métricos. Este modelo, además de brindar un marco formal, es independiente del dominio de la aplicación. Un espacio métrico está compuesto por un *universo* $\mathcal{U}$ de objetos y una *función de distancia* $d : \mathcal{U} \times \mathcal{U} \longrightarrow \mathbb{R}^+$, la cual cumple con las propiedades de una métrica. Sobre una *base de datos* $\mathcal{S} \subseteq \mathcal{U}$, se pueden considerar dos tipos básicos de búsqueda por similitud: la *búsqueda por rango* y la *búsqueda de los k vecinos más cercanos*. La función de distancia permite medir el mínimo esfuerzo (costo) necesario para transformar un objeto en otro. Sin embargo, para algunos de los tipos de datos no estructurados, el cálculo de la distancia puede ser muy costoso. Por lo tanto, un objetivo de todo método de acceso es ahorrar cálculos de distancia y en su gran mayoría para lograrlo aprovechan que la función de distancia satisface la desigualdad triangular.

Si se considera que la base de datos $\mathcal{S}$ posee n objetos, cualquier consulta puede responderse de manera trivial realizando n evaluaciones de distancia. Sin embargo, en la mayoría de las aplicaciones sobre grandes volúmenes de datos y siendo las distancias costosas de computar (por ej.: comparación de huellas digitales), no es factible aplicar la solución trivial. Por lo tanto, si el objetivo es responder consultas realizando la menor cantidad de cálculos de distancia posibles, se debe construir un índice a través del preprocesamiento de la base de datos. En algunos casos particulares, es probable que la base de datos, el índice, o ambos, no puedan almacenarse en memoria principal y deban hacer uso de niveles más bajos de la jerarquía de memorias, como es la memoria secundaria. Pero ello acarrea altos costos en las operaciones de E/S. Por lo tanto, para lograr eficiencia, se debe minimizar el número de operaciones de E/S, considerar el nivel de la jerarquía de memorias sobre la que se trabaja y en algunos casos admitir respuestas no exactas; utilizando, cuando sea posible, técnicas de computación paralela.

En este contexto, se considera como objetivo principal obtener herramientas de recuperación de información para procesar conjuntos masivos de datos, desarrollando nuevas técnicas y aplicaciones que soporten la interacción con el usuario, diseñando índices que permitan la manipulación eficiente de grandes volúmenes de datos no estructurados y faciliten la realización de diferentes tipos de consultas. De esta manera, se espera contribuir al desarrollo de aplicaciones reales para problemas de big data.

3. Líneas de Investigación

Dado que en esta investigación se pretende contribuir a distintos aspectos de los sistemas de recuperación de información sobre grandes volúmenes de datos no estructurados, se ha considerado el diseño de nuevas medidas de similitud, de nuevos índices y la resolución de distintas consultas sobre estos tipos de bases de datos y cómo lograr eficiencia y escalabilidad para grandes volúmenes de datos.

Índices

Los índices métricos resultan apropiados para realizar búsquedas sobre bases de datos conteniendo datos no estructurados [5]. Todos ellos aprovechan que la función de distancia debe cumplir la propiedad de desigualdad triangular para ahorrar algunos cálculos de distancia y, de esta manera, ahorrar tiempo. 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, almacenando un radio de cobertura que determina la zona de cada centro.

Los aspectos que se consideran de interés al diseñar índices incluyen: dinamismo, en qué nivel de la jerarquía de memorias deben almacenarse, si pueden aplicar técnicas de computación de alto desempeño para mejorar los tiempos de respuesta, si deben proporcionar una respuesta exacta o basta con una respuesta aproximada y la dimensionalidad del espacio métrico considerado.

Como los conjuntos de datos masivos de interés en este caso son aquellos que contienen datos no estructurados, los volúmenes de información con los que se debe trabajar (por ejemplo, millones de imágenes en la Web) hacen 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, que efectúen 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. En particular, se ha diseñado e implementado una versión paralela del *Conjunto Dinámico de Clusters* (DSC) [13]. Este índice, basado en la *Lista de Clusters* (LC) [4], 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 Aproxiación Espacial* [11], es un índice estático que no necesita sintonizar ningún parámetro y es muy eficiente gracias a definir una partiición de hiperplanos con muy buenas características [3]. 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 particiones 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 primero. 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 permite calcular, logrando una respuesta aproximada a la consulta por similitud con poco costo.

Por otra parte, se está estudiando cómo aprovechar los índices para búsquedas por similitud sobre conjuntos 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.

Nuevas Medidas de Similitud

Existem numerosos algoritmos que permiten resolver eficientemente las búsquedas cuando se consideran espacios de baja dimensión. Sin embargo, su desempeño empeora a medida que la dimensionalidad intrínseca del espacio crece [5, 12]. Más aún, en bases de datos cuya dimensionalidad intrínseca es alta, el desempeño se puede degradar de tal forma que sea equivalente al de haber realizado una búsqueda exhaustiva sobre la base de datos completa [5]. Por lo tanto, el desafío se encuentra en esa clase

de base de datos, donde el histograma de las distancias entre los objetos es muy concentrado; es decir, donde todas las distancias entre pares de objetos de la base de datos son casi iguales.

Una aproximación práctica sobre este tipo de base de datos es resignarse a no obtener la respuesta por similitud exacta para las consultas. En su lugar, es posible conformarse con respuestas aproximadas; lo cual significa que se admite que se pierdan algunos objetos relevantes desde el conjunto de objetos de la respuesta o que se reporten en dicho conjunto algunos elementos que no sean relevantes [6]. Así, el objetivo es diseñar métodos eficientes cuya calidad de la respuesta esté dentro de ciertos límites.

En [1], se presenta un nuevo método aproximado para búsquedas por similitud, cuyo desempeño es insuperable en bases de datos de alta dimensión [10]. Sin embargo, se puede mejorar aún más su desempeño si se considera una medida diferente de similitud entre permutaciones [7].

La medida más utilizada de distancia entre permutaciones es *Spearman Footrule*, la cual se define como $S_F(\Pi_u, \Pi_q) = \sum_{1 \leq i \leq m} |\Pi_u^{-1}(i) - \Pi_q^{-1}(i)|$, donde Π_u y Π_q son las permutaciones de u y q.

Las nuevas medidas de similitud entre permutaciones que en particular, como se demuestra en [7], no cumplen con ser una métrica sino una semimétrica, se basan en particionar las permutaciones en trozos y utilizar en los trozos significativos (el inicial y el final) un factor que permita amplificar las grandes diferencias de posiciones de los permutantes y descartar el trozo medio menos significativo.

DBMS 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 [15].

Existen diferentes variantes para el join por similitud, dependiendo del criterio de similitud Φ utilizado, pero ellas tienen en común que se aplican entre dos bases de datos A y B, ambas 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 Φ 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 ($A \underset{\Phi}{\bowtie} B$) 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 $\Phi(x, y)$ es verdadero (se satisface el criterio de similitud Φ 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 [14], por lo tanto 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 [2]. Por otro lado, no se dispone de este tipo de consultas para todo tipo de datos métricos. Así, es importante proveer un DBMS para todos los posibles datos métricos y sus operaciones [9].

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 [16]; 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.

4. Resultados

Se ha publicado en [7] una familia de medidas de similitud para permutaciones que permiten mejorar el desempeño de los algoritmos basados en permutaciones [1]. Además, se ha publicado en [8] una nueva estructura para búsquedas aproximadas, especialmente diseñada para trabajar con grandes volúmenes

de datos en memoria secundaria.

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 considere opciones de respuesta aproximada, como también la posibilidad de diversificación de respuestas para 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 desarrollo un trabajo final de la Ingeniería en Computación.

Referencias

[1] E. Chávez, K. Figueroa, and G. Navarro. Effective proximity retrieval by ordering permutations. *IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI)*, 30(9):1647–1658, 2009.

[2] E. Chávez, V. Ludueña, and N. Reyes. Revisiting the VP-forest: Unbalance to improve the performance. In *Proc. de las JCC08*, page 26, 2008.

[3] E. Chávez, V. Ludueña, N. Reyes, and P. Roggero. Faster proximity searching with the distal sat. *Information Systems*, 59:15 – 47, 2016.

[4] E. Chávez and G. Navarro. A compact space decomposition for effective metric indexing. *Pattern Recognition Letters*, 26(9):1363–1376, 2005.

[5] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroquín. Searching in metric spaces. *ACM*, 33(3):273–321, September 2001.

[6] P. Ciaccia and M. Patella. Approximate and probabilistic methods. *SIGSPATIAL Special*, 2(2):16–19, 2010.

[7] K. Figueroa, R. Paredes, and N. Reyes. New permutation dissimilarity measures for proximity searching. In *Similarity Search and Applications - 11th International Conference, SISAP 2018*, volume 11223 of *LCNS*, pages 122–133. Springer, 2018.

[8] K. Figueroa, N. Reyes, A. Camarena-Ibarrola, and L. Valero-Elizondo. Improving the list of clustered permutation on metric spaces for similarity searching on secondary memory. In *Pattern Recognition*, pages 82–92, Cham, 2018. Springer.

[9] F. Kasián and N. Reyes. Búsquedas por similitud en PostgreSQL. In *Actas del XVIII CACiC*, pages 1098–1107, Oct. 2012.

[10] B. Naidan, L. Boytsov, and E. Nyberg. Permutation search methods are efficient, yet faster search is possible. *Proc. VLDB Endow.*, 8(12):1618–1629, August 2015.

[11] G. Navarro. Searching in metric spaces by spatial approximation. *VLDBJ*, 11(1):28–46, 2002.

[12] G. Navarro, R. Paredes, N. Reyes, and C. Bustos. An empirical evaluation of intrinsic dimension estimators. *Information Systems*, 64:206 – 218, 2017.

[13] G. Navarro and N. Reyes. New dynamic metric indices for secondary memory. *Information Systems*, 59:48 – 78, 2016.

[14] 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.

[15] 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.

[16] L. Santos, L. Carvalho, W. Oliveira, A. Traina, and C. Jr. Traina. Diversity in similarity joins. In *Similarity Search and Applications*, volume 9371 of *LNCS*, pages 42–53. Springer, 2015.

Selección de características mediante la combinación de métodos para evaluar la precisión de clasificación en un conjunto de datos de implantes dentales

Nancy B. Ganz[1,*], Facundo A. Domínguez[2], Alicia E. Ares[1] y Horacio D. Kuna[2]
[1] Instituto de Materiales de Misiones, IMAM (CONICET-UNaM). Félix de Azara 1552, N3300LQH, Posadas, Misiones, Argentina.
[2] Departamento de Computación. Facultad de Ciencias Exactas, Químicas y Naturales. Universidad Nacional de Misiones. Félix de Azara 1552, N3300LQH, Posadas, Misiones, Argentina.
*E-mail: nancy.bea.ganz@gamil.com, facundokpo04@gmail.com, a.e.ares@gmail.com, hdkuna@gmail.com

RESUMEN

La selección de características es una técnica de preprocesamiento que permite encontrar un conjunto reducido de características, el cual concentra la información más sustancial del conjunto de datos. En este trabajo, se propone un procedimiento para la selección de las características más relevantes de un conjunto de datos de implantes dentales, de la Provincia de Misiones, Argentina. Se basa en la combinación de los métodos Information Gain, Gain Ratio, Random Forest importance, Relief y Chi Squared con el fin de predecir la clase minoritaria (Fracaso). El rendimiento del procedimiento propuesto se evaluó no sólo mediante la precisión de clasificación, en cuanto a las medidas de rendimiento tnr y bac de los clasificadores SVM rbf y Naive Bayes con validación cruzada, sino que también en base a la cantidad de características seleccionadas. Se observó que el procedimiento propuesto seleccionó la cantidad de características más adecuado para el estudio de caso y mejoró la precisión en la clasificación para la clase minoritaria.

Palabras Clave: *Ganancia de Información, Métodos, Selección de Características, Fracaso, Implantes Dentales.*

CONTEXTO

Esta línea de investigación se lleva a cabo dentro del Programa de Materiales y Fisicoquímica (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, bajo el nombre de *"Aplicación de la Minería de Datos para la Selección de Biomateriales"*. Está financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) a través de una "Beca Interna Doctoral" otorgada por Resolución D Nº 4869. Además, cuenta con el auspicio del Colegio de Odontólogos de la Provincia de Misiones.

1. INTRODUCCIÓN

Este trabajo presenta un enfoque para la selección de características basado en la combinación de cinco métodos, para encontrar el subconjunto de características más relevante, evaluando la calidad de precisión en la clasificación y el número de características seleccionadas por cada método. La experimentación consistió en obtener los valores de importancia de cada característica en función de la integración de los métodos de selección de características: Information Gain[1], Gain Ratio[2], Random Forest importance[3], Relief[4] y Chi Squared[5]. Los pasos realizados fueron básicamente tres: generación del subconjunto de características, obtención de las medidas de rendimiento y apreciación de esas medidas para contrastar con el procedimiento propuesto. Para el propósito de este trabajo, se utilizó un conjunto de datos de historias clínicas de pacientes que se han sometido a procesos quirúrgicos de colocación de implantes dentales en la Provincia de Misiones, Argentina. Este conjunto de datos se encuentra representado a través de 4 dimensiones: datos del paciente

(antecedentes y condiciones médicas de los pacientes a la hora de la intervención), datos del implante (características del implante utilizado por el especialista implantólogo), datos de la fase quirúrgica (procesamiento de intervención quirúrgica y mejoramiento del lecho óseo del paciente) y datos del seguimiento postoperatorio (resultado del proceso de colocación del implante, es decir si el proceso de oseointegración implante/tejido-óseo tuvo éxito o fracasó). Para lograr este conjunto de datos se siguió un proceso, empleando la metodología CRISP-DM. Donde, en la primera fase correspondiente a la comprensión del problema, se evaluó la necesidad de la selección de características en conjuntos de datos desbalanceados. En la segunda fase se exploraron y estudiaron los datos, dimensiones, peculiaridades y verificación de calidad de los mismos, junto a los expertos. En la tercera fase, se procedió a la preparación y limpieza de los datos. Luego, se planteó un procedimiento para la selección de características, se probó y evaluó. Los trabajos[6]–[12] utilizan estos métodos de selección de características para reducir la dimensionalidad de los datos. Así mismo, utilizan comúnmente como clasificadores a Naive Bayes[13] y SVM[14] con validación cruzada. Y a la hora de medir la precisión de estos clasificadores, utilizan medidas de rendimiento como: matriz de confusión[15], accuracy, precisión y curva roc[16]. Es por esto que se ha tenido en cuenta estas particularidades para el diseño y validación del procedimiento propuesto, además de propiciar la combinación de los cinco métodos de selección de características propuestos para ofrecer la máxima precisión posible y no sesgar la decisión sobre los resultados de uno solo.

Debido a que la experimentación del procedimiento se realizó sobre un conjunto de datos desbalanceado, se buscó validar específicamente las métricas de rendimiento tnr y bac, debido a que son las medidas de referencia al tratarse de un conjunto de datos desbalanceado[17], [18].

Conjuntamente, se incorporó el análisis del método de selección de características que se extiende del algoritmo Random Forest, el cual se denomina Random Forest importance.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación propone estudiar, evaluar y aplicar distintas técnicas de ciencia de datos, para detectar los factores de fracaso de una base de datos de implantes dentales de la Provincia de Misiones. A través de la presente línea de investigación, se logró confeccionar el conjunto de datos necesario para proceder al diseño de un procedimiento, que a través de una metodología hibrida, permita identificar los factores de fracaso. Así mismo, se reconoce como objetivo de interés, el estudio de los tipos de tratamientos de superficie de los implantes utilizados en la zona, para incorporar como característica a evaluar.

3. RESULTADOS OBTENIDOS

Entre los resultados obtenidos, destacamos la creación de un procedimiento que consistió en:

Paso 1. Leer el conjunto de datos.

Paso 2. Seleccionar la característica objetivo para la predicción.

Paso 3. Obtener los subconjuntos de características de los métodos:

Information Gain (IG):

$$IG(X) = H(Y) - H(Y|X) = H(X) - H(X|Y)$$

Gain Ratio (GR):

$$GR = \frac{IG}{H(X)}$$

Random Forest importance (RFI):

$$RFI(x_m) = \frac{1}{N_T}\sum_{T}\sum_{t\epsilon T:\upsilon(S_t)=X_m} \mathrm{p(t)}\Delta i(\mathrm{s_t}, \mathrm{t})$$

Relief (R):

$$R_f = P\left(\frac{\text{different value of f}}{\text{class different}}\right) - P\left(\frac{\text{different value of f}}{\text{same class}}\right)$$

Chi Squared (ChiS):

$$X^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{(O_{ij}-E_{ij})^2}{E_{ij}}$$

Paso 4. Confeccionar una matriz que concentre el valor de importancia obtenido por los distintos métodos para cada característica.

Paso 5. Normalizar los valores, debido a que los métodos empleados se desempeñan con rangos diferentes, este paso es necesario y fundamental para lograr un valor medio para cada característica. Se utilizó la función *normalize*, esta permite normalizar valores en base al método mínimo-máximo. La normalización mínimo-máximo regulariza las características en un rango[19]. Dado min_A y max_A valores mínimo y máximo de una característica A.

La normalización mínimo-máximo mapea un valor v_i de A para v_i' en el rango $[new_min_A, new_max_A]$ mediante $v_i' = \frac{v_i - min_A}{max_A - min_A}(new_max_A - new_min_A) + new_min_A$.

Se utilizó este criterio de normalización debido a que este permite preservar todas las relaciones de los valores de los datos originales, es decir no introduce ningún sesgo potencial en los datos. Además, se encuentra demostrado que tiene mejor rendimiento en la clasificación[20], [21]. El rango empleado fue [0,1].

Paso 6. Transponer la matriz.

Paso 7. Obtener la mediana de cada característica en función de los valores arrojados por los distintos métodos. Se empleó la mediana como medida de tendencia central debido a que los valores de importancia arrojados por los distintos métodos no seguían una distribución normal. En el caso de que estos valores sigan una distribución normal aplicar la media[19].

Paso 8. Ordenar la matriz en forma decreciente.

Paso 9. Obtener un umbral, a través de validación cruzada de 10 iteraciones. De esta manera se logró hallar el umbral más óptimo para las características cuyos valores de importancia resultaron del paso 7.

Paso 10. Seleccionar las características que cumplan con la condición de ser igual o mayor al umbral obtenido en el paso 9.

Paso 11. Obtener un umbral óptimo para cada uno de los 5 métodos, esto se realizó a través de una calibración del método sobre el conjunto de datos con un clasificador Naive Bayes y una validación cruzada de 10 veces. Es necesario aclarar, que el umbral conjuntamente fue ajustado en función del conocimiento que se tiene sobre los datos.

Paso 12. Seleccionar las características para cada uno de los 5 métodos que cumplan con la condición de ser igual o mayor al umbral obtenido en el paso 11.

Paso 13. Aplicar los clasificadores SVM $f(X) = W^T\emptyset(X)$ con núcleo rbf $K(x_i, x_j) = \emptyset\,(x_i).\emptyset(x_j)$ y Naive Bayes $f_i(X) = \prod_{j=1}^{N} P(x_j|c_i)\,P(c_i)$ con una validación cruzada de 100 iteraciones sobre los 5 conjuntos de características obtenidos en el paso 12 y a la matriz normalizada lograda en el paso 10.

Paso 14. Obtener las medidas de rendimiento: tpr, fpr, tnr, fnr, bac, auc y acc.

Paso 15. Validar las medidas tnr y bac.

Cabe aclarar, que tanto a los métodos IG, GR, RFI, R y ChiS, así como a los clasificadores, se los consideró a cada uno con el mismo peso. Debido a que la finalidad fue optimizar la cantidad y conocer que características seleccionaba cada método, para posteriormente corroborar su rendimiento individual con los clasificadores.

Para la funcionalidad del procedimiento propuesto se utilizó un conjunto de datos de historias clínicas de implantes dentales, con 1.050 filas, 31 características y un atributo clase binario: EXITO (977 casos) y FRACASO (73 casos).

La implementación se efectuó sobre la herramienta RStudio versión 3 de GNU con licencia AGPL (Affero General Public License), cuyo software es gratuito y de código abierto[22]. Conjuntamente, se utilizó la librería MLR para disponer de los métodos de selección de características. Este es un paquete muy completo, el cual proporciona métodos supervisados como clasificación y regresión. Así mismo, provee métodos no supervisados como agrupación, junto con métodos de evaluación y optimización[23]. Al mismo tiempo, el paquete MLR utiliza los paquetes FSelector[24] y randomForest[25] para obtener los algoritmos de selección de características.

La Tabla 1 muestra la cantidad de características seleccionadas por los métodos utilizados y el procedimiento propuesto (PP), así como los resultados arrojados con los clasificadores.
En cuanto a los valores arrojados de la clasificación, se observó que comparando con los métodos IG, GR, RFI, R y ChiS, PP logró la mejor tasa de verdaderos negativos (tnr) con el clasificador Naive Bayes y una medida bac de 53%.

Tabla 1 – Resultados obtenidos de aplicar los métodos IG, GR, RFI, R, ChiS y el procedimiento propuesto con los clasificadores SVM rbf y Naive Bayes.

Métodos	Cant.	tpr	fpr	tnr	fnr	bac	auc	acc
IG-SVM	10	0.997	0.989	0.011	0.003	0.504	0.698	0.929
IG-NB		0.982	0.949	0.051	0.018	0.517	0.720	0.917
GR-SVM	11	0.999	1.000	0.000	0.001	0.500	0.563	0.930
GR-NB		0.992	0.937	0.063	0.008	0.527	0.726	0.927
RFI-SVM	22	0.998	0.938	0.062	0.002	0.530	0.680	0.933
RFI-NB		0.974	0.958	0.042	0.026	0.508	0.694	0.909
R-SVM	18	0.998	0.978	0.022	0.002	0.510	0.676	0.930
R-NB		0.981	0.916	0.084	0.019	0.533	0.697	0.919
ChiS-SVM	16	0.998	0.999	0.001	0.002	0.500	0.678	0.929
Chis-NB		0.975	0.918	0.083	0.025	0.529	0.734	0.913
PP-SVM	17	0.996	0.974	0.026	0.004	0.511	0.710	0.929
PP-NB		0.975	0.913	**0.087**	0.025	**0.531**	0.732	0.913

Basándonos en los resultados, podemos afirmar que los algoritmos de selección de características utilizados, si se los emplea de forma individual, pueden ocasionar una incorrecta selección de subconjuntos de características. Una adecuada combinación de estos métodos, puede producir subconjuntos de características más efectivas para la clasificación, como ocurre en este caso.
También hemos contemplado que, los métodos de selección de características IG y GR lograron reducir en ocasiones más la cantidad de características que los métodos RFI, R, ChiS y PP, pero su precisión de clasificación con los clasificadores SVM rbf y NB no fueron tan buenos. Además, PP seleccionó la cantidad más adecuada de características, siendo éstas las más efectivas en la clasificación.
Se observó, que el clasificador NB fue superior respecto a SVM, debido a que logró mejores rendimientos, y sobre todo al tratarse de un conjunto de datos desbalanceado. Conjuntamente, se valoró que SVM requiere de mayor calibración y necesita conocer a priori los datos, para mejorar su rendimiento y obtener resultados comparables.
Por lo tanto, podemos afirmar que PP es útil cuando se requiere definir aquellas características que se encuentran en un border line, además combina varios métodos y no sesga la decisión sobre los resultados de una sola técnica. También, en base a la opinión de los especialistas, se pudo validar que los factores detectados son los que mayor influencia ejercen sobre el proceso de oseointegración.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto es parte de las líneas de investigación del "Programa de Materiales y Fisicoquímica" de la FCEQyN – UNaM, con cuatro integrantes: un investigador independiente del CONICET, un docente categoría I perteneciente al Depto. de Informática de la FCEQyN – UNaM, un

doctorando y un tesista de grado de la carreara de Licenciatura en Sistemas de Información.

5. BIBLIOGRAFÍA

[1] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.

[2] J. R. Quinlan, "Induction of Decision Trees," *Mach. Learn.*, vol. 1, no. 1, pp. 81–106, 1986.

[3] L. Breiman, "Random Forest," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[4] K. Kira and L. A. Rendell, "A Practical Approach to Feature Selection," *Mach. Learn. Proc.*, pp. 249–256, 1992.

[5] K. Pearson, "On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling," *London, Edinburgh, Dublin Philos. Mag. J. Sci.*, vol. 50, no. 302, pp. 157–175, 1900.

[6] A. Chaudhary, S. Kolhe, and Rajkamal, "Performance Evaluation of feature selection methods for Mobile devices," *J. Eng. Res. Appl.*, vol. 3, no. 6, pp. 587–594, 2013.

[7] Z. Karimi, M. Mansour, R. Kashani, and A. Harounabadi, "Feature Ranking in Intrusion Detection Dataset using Combination of Filtering Methods," *Int. J. Comput. Appl.*, vol. 78, no. 4, pp. 21–27, 2013.

[8] L. Gao, M. Ye, X. Lu, and D. Huang, "Hybrid Method Based on Information Gain and Support Vector Machine for Gene Selection in Cancer Classification," *Genomics, Proteomics Bioinforma.*, vol. 15, no. 6, pp. 389–395, 2017.

[9] T. Z. Phyu and N. N. Oo, "Performance Comparison of Feature Selection Methods," *MATEC Web Conf.*, vol. 42, p. 06002, 2016.

[10] H. Dag, K. E. Sayin, I. Yenidogan, S. Albayrak, and C. Acar, "Comparison of Feature Selection Algorithms for Medical Data," *Int. Symp. Innov. Intell. Syst. Appl.*, pp. 1–5, 2012.

[11] M. Peker, A. Arslan, B. Sen, F. V. Celebi, and A. But, "A novel hybrid method for determining the depth of anesthesia level: Combining ReliefF feature selection and random forest algorithm (ReliefF+RF)," *Int. Symp. Innov. Intell. Syst. Appl.*, 2015.

[12] W. Li, Y. Li, J. Chen, and C. Hou, "Product functional information based automatic patent classification: Method and experimental studies," *Inf. Syst.*, vol. 67, pp. 71–82, Jul. 2017.

[13] N. Friedman, D. Geiger, and M. Goldszmit, "Bayesian Network Classifiers," *Mach. Learn.*, vol. 29, pp. 131–163, 1997.

[14] C. Cortes and V. Vapnik, "Support-Vector Networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.

[15] R. Susmaga, "Confusion Matrix Visualization," in *Intelligent Information Processing and Web Mining*, Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 107–116, 2004.

[16] M. Sokolova and G. Lapalme, "A systematic analysis of performance measures for classification tasks," *Inf. Process. Manag.*, vol. 45, no. 4, pp. 427–437, 2009.

[17] Q. Wei and R. L. Dunbrack, "The Role of Balanced Training and Testing Data Sets for Binary Classifiers in Bioinformatics," *PLoS One*, vol. 8, no. 7, 2013.

[18] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, 2009.

[19] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*. 2012.

[20] W. Li and Z. Liu, "A method of SVM with normalization in intrusion detection," *Procedia Environ. Sci.*, vol. 11, no. PART A, pp. 256–262, 2011.

[21] S. Jain, S. Shukla, and R. Wadhvani, "Dynamic selection of normalization techniques using data complexity measures," *Expert Syst. Appl.*, vol. 106, pp. 252–262, 2018.

[22] "RStudio - RStudio." [Online]. Available: https://www.rstudio.com/products/rstudio/. [Accessed: 26-Feb-2019].

[23] B. Bischl *et al.*, "mlr: Machine Learning in R," *J. Mach. Learn. Res.*, vol. 17, pp. 1–5, 2016.

[24] P. Romanski and L. Kotthoff, "Package 'FSelector,'" *Repository CRAN*. 2018.

[25] L. Breiman and A. Cutler, "Package 'randomForest,'" *Repository CRAN*. 2018.

Técnicas de Análisis de Sentimientos Aplicadas a la Extracción de Opiniones en el Lenguaje Español

Rosenbrock Germán, Trossero Sebastián, Goette Pablo, Llorente María Emilia, Pascal Andrés, Cían Damián

Laboratorio de Análisis, Procesamiento, Almacenamiento y Control de Datos
Facultad de Ciencia y Tecnología
Universidad Autónoma de Entre Ríos

german009@gmail.com, sebastian@lambdasi.com.ar, pgoette@hexacta.com, maria.fcyt@gmail.com, andrespascal2003@yahoo.com.ar, damiancian@gmail.com

RESUMEN

Actualmente existe una gran cantidad de datos textuales disponibles, principalmente en Internet, que crece día a día. El texto es el tipo de dato más utilizado en la web, ya que es fácil de publicar y generar. Lo complejo es obtener información a partir de los mismos en forma automática, y manualmente es sumamente costoso.

La información textual puede dividirse en dos tipos principales: hechos y opiniones. Mientras que los hechos son objetivos, las opiniones representan los sentimientos de cada autor. La Minería de Opinión o Análisis de Sentimientos estudia la extracción de información a partir de datos subjetivos y es relativamente reciente.

Desde hace ya varios años existen sitios web donde los usuarios pueden expresar sus opiniones respecto a diversos temas, por ejemplo, nuevos productos o servicios, imagen empresarial, propuestas de leyes, etc.

Este proyecto propone analizar distintas técnicas de Análisis de Sentimiento aplicadas a opiniones expresadas en el lenguaje Español, evaluar sus resultados para distintos casos reales, y realizar mejoras a las mismas.

CONTEXTO

Este trabajo se encuentra en el marco del Proyecto de Investigación y Desarrollo de Inserción (PIDIN) denominado "Evaluación y Mejora de Técnicas de Minería de Opinión / Análisis de Sentimientos", de la Facultad de Ciencia y Tecnología de Oro Verde de la Universidad Autónoma de Entre Ríos.

Su ejecución forma parte de las actividades del Laboratorio de Análisis, Procesamiento, Almacenamiento y Control de Datos (LAPACDA) perteneciente a la mencionada facultad.

1. INTRODUCCION

En un proceso de toma de decisiones, es fundamental contar con la información oportuna, confiable e íntegra que permita un análisis real de la situación en estudio. En ciertos casos, los datos de origen son opiniones personales. En forma previa a la Web 2.0, su importancia no era alta debido a la falta de grandes cantidades de textos que registren opiniones. Con la disponibilidad masiva de este tipo de información, surgen nuevas oportunidades y desafíos en la búsqueda, comprensión e interpretación de la misma. Sin embargo, la búsqueda en

estos sitios y la posterior valoración de las opiniones en forma manual es un trabajo intenso y costoso, por lo tanto, es necesario contar con sistemas que automaticen este proceso.

El Análisis de Sentimientos o Minería de Opiniones estudia la interpretación automática de opiniones y sentimientos expresados mediante el lenguaje natural.

Es utilizada por ejemplo, por algunas organizaciones para el análisis de su imagen. Además, la literatura muestra varios otros tipos de aplicaciones, incluyendo: valoración de películas [1], opiniones sobre deportes [2], turismo [3, 4], política [5], educación [6], salud [7], finanzas [8] y automóviles [9].

Este trabajo propone aplicar las técnicas Máquinas de Vectores de Soporte (SVM) [2, 10, 11], Clasificador Bayesiano (Naive-Bayes) [12, 13, 14], Redes Neuronales Profundas (DNN) [15, 16], Máxima Entropía [17] y K-Vecinos más Cercanos [18], a conjuntos de opiniones expresadas en el lenguaje español, para luego evaluar sus comportamientos y proponer mejoras.

Los casos de estudio son tres:

- www.cinesargentinos.com.ar (revisión de películas)
- www.booking.com/hotel/ar (comentarios sobre hoteles, versión español)
- www.mercadolibre.com.ar/ (opiniones sobre productos)

La selección de estos sitios se realizó teniendo como criterios la disponibilidad de los datos, la cantidad de opiniones, el nivel de informalidad en el uso del lenguaje, la disponibilidad de una valoración ya registrada para cada opinión, y la existencia de distintos aspectos a evaluar por cada opinión.

2. LINEAS DE INVESTIGACION Y DESARROLLO

La línea principal de investigación es la Minería de Opiniones, para la cual se utilizan técnicas de Inteligencia Artificial; en particular, Procesamiento del Lenguaje Natural (PLN).

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo general de este trabajo consiste en evaluar la aplicación de análisis de sentimientos para la obtención de calificaciones cuantitativas a partir de valoraciones textuales cualitativas de reseñas generadas por usuarios de sitios web en el idioma español/castellano, ya que la mayoría de los estudios actuales están hechos sobre el lenguaje Inglés.

Los objetivos específicos son:

- Evaluación de técnicas de análisis de sentimientos aplicadas a opiniones escritas en español/castellano. Las técnicas principales a ser evaluadas son: Naive Bayes, SVM, Entropía Máxima y combinaciones de las anteriores con el uso de diccionarios/léxicos. Además se agregará al menos una más, a seleccionar entre Random Forest, Redes Neuronales Profundas y K-NN, teniendo como criterio su grado de complementación con las anteriores (en el leguaje español), para definir enfoques híbridos.
- Establecer cuáles obtienen mejores resultados para los distintos casos propuestos.
- Realizar mejoras a dichas técnicas.

El aporte principal de este trabajo es el diseño de una técnica híbrida que permita obtener mejores resultados que las técnicas nombradas anteriormente, particularmente en su aplicación a opiniones escritas en español/castellano.

Actualmente hemos extraído datos del sitio www.cinesargentinos.com.ar y conformado una base de datos con 52.307

opiniones sobre diversas películas. Luego utilizamos técnicas de PLN (análisis léxico, eliminación de stopwords, stemming) para extraer las palabras de mayor relevancia para el análisis. Posteriormente entrenamos un modelo usando Naive-Bayes y estamos realizando ajustes y validaciones para asegurar la calidad de los resultados.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo está formado en su totalidad por miembros del Laboratorio de Análisis, Procesamiento, Almacenamiento y Control de Datos (LAPACDA) de la Facultad de Ciencia y Tecnología de la UADER de la sede de Concepción del Uruguay y de Oro Verde. Participan docentes de la carrera Lic. en Sistemas de Información de la UADER (dos de ellos realizando una Maestría en Minería de Datos), y alumnos avanzados de la carrera de grado.

5. BIBLIOGRAFIA

[1] Kuat Yessenov. Sentiment Analysis of Movie Review Comments. 2009.

[2] N. LI and D. D. W. Using text mining and sentiment analysis for online forums hotspot detection and forecast. Decision Support Systems, vol. 48, nº 2, pp. 354 - 368, 2010.

[3] L. C. Fiol, J. S. García, M. M. T. Miguel and S. F. Coll, «La importancia de las comunidades virtuales para el análisis del valor de marca. El caso de TripAdvisor en Hong Kong y París,» Papers de turisme, nº 52, pp. 89-115, 2012.

[4] C. Henriquez, J. Guzmán and D. Salcedo. Minería de Opiniones basado en la adaptación al español de ANEW sobre opiniones acerca de hoteles. Procesamiento del Lenguaje Natural, vol. 56, pp. 25-32., 2016.

[5] S. Rill, D. Reinel, J. Scheidt and R. V. Zicari. PoliTwi: Early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis. Knowledge-Based Systems, vol. 69, pp. 24-33, 2014.

[6] A. Ortigosa, J. M. Martín and R. M. Carro. Sentiment analysis in Facebook and its application to e-learning. Computers in Human Behavior, vol. 31, pp. 527-541, 2014.

[7] F. Greaves, D. Ramirez-Cano, C. Millett, A. Darzi and L. Donaldson. Use of Sentiment Analysis for Capturing Patient Experience From Free-Text Comments Posted Online. Journal of medical Internet research, vol. 15, nº 11, 2013.

[8] X. Dong, Q. Zou and Y. Guan. Set-Similarity joins based semi-supervised sentiment analysis. Neural Information Processing. Springer Berlin Heidelberg, 2012., from Neural Information Processing, Springer Berlin Heidelberg, 2012, pp. 176-183.

[9] P. D. Turney. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. Proceedings of the 40th annual meeting on association for computational linguistics, Stroudsburg, PA, USA, 2002.

[10] A. Abbasi, H. Chen and A. Salem. Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums. ACM Transactions on Information Systems (TOIS), vol. 26, nº 3, p. 12, 2008.

[11] F. Pla and L.-F. Hurtado. Sentiment Analysis in Twitter for Spanish. Natural Language Processing and Information Systems, pp. 208-213, 2014.

[12] A. Lazaridou, I. Titov and C. Sporleder. A Bayesian Model for Joint Unsupervised Induction of Sentiment, Aspect and Discourse Representations.. ACL, vol. 1, pp. 1630 -1639, 2013.

[13] A. Pak and P. Paroubek. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. LREC, vol. 10, pp. 1320 -1326, 2010.

[14] J. Ortigosa-Hernández, J. D. Rodríguez, L. Alzate, M. Lucania, I. Inza and J. A. Lozano. Approaching Sentiment Analysis by using semi-supervised learning of multi-dimensional classifiers. Neurocomputing, vol. 92, pp. 98-115, 2012.

[15] M. Ghiassi, J. Skinner and D. Zimbra. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. vol. 40, nº 16, pp. 6266-6282, 2013.

[16] M. Anjaria and R. M. R. Guddeti. A novel sentiment analysis of social networks using supervised learning. Social Network Analysis and Mining, vol. 4, nº 1, pp. 1-15, 2014.

[17] Pang, Bo & Lee, Lillian & Vaithyanathan, Shivakumar. (2002). Thumbs up? Sentiment Classification Using Machine Learning Techniques. EMNLP. 10. . 10.3115/1118693.1118704.

[18] Jędrzejewski, K. & Zamorski, M. (2013). Performance of K-Nearest Neighbors Algorithm in Opinion Classification. Foundations of Computing and Decision Sciences, 38(2), pp. 97-110. 2017.

Text Mining aplicado a la Base de Datos del Servicio del 911 de la Provincia de Jujuy.

José Federico Medrano, Mario Alberto Tejerina, César Alejandro Castillo, Juan Carlos Rodríguez

jfmedrano@fi.unju.edu.ar, mariotejerina@gmail.com, ce_al_castillo@yahoo.com.ar, jcrodriguez@cegin.com.ar

VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

RESUMEN

La minería de texto o *text mining* la englobamos dentro de las técnicas y modelos de minería de datos o *data mining*, que es el análisis matemático para deducir patrones y tendencias que existen en los datos, patrones que no pueden detectarse mediante una exploración tradicional de los datos porque las relaciones son demasiado complejas o por el volumen de datos que se maneja. A diferencia del *data mining*, el *text mining* trabaja sobre datos no estructurados en grandes colecciones de datos.

La base de datos del Servicio del 911 de la provincia de Jujuy presenta un panorama muy interesante, ya que si bien la información que maneja dicho servicio es almacenada en una base de datos estructurada, la tabla principal que almacena los incidentes que son cargados por los tele-operadores posee un campo llamado descripción donde allí se vuelca la información más importante y los detalles del llamado de emergencia, información que no puede almacenarse en los campos del formulario de carga.

Este campo ofrece una alternativa muy interesante de exploración pues el contenido no estructurado de dicho campo no es analizado ni tenido en cuenta por los informes tradicionales que utilizan en la actualidad.

Palabras clave: *Text Mining; Clasificación de Información; Clustering; Data Mining; PLN*

CONTEXTO

La línea de investigación aquí presentada se encuentra inserta en el proyecto: *"Text Mining aplicado a la Base de Datos del Servicio del 911 de la Provincia de Jujuy."*, ejecutado a partir de enero de 2019 con una duración de 1 año. 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 Facultad de Ingeniería de la Universidad Nacional de Jujuy (Resolución FI N° 642/18)

1. INTRODUCCIÓN

Las cantidades de texto que se genera todos los días están aumentando drásticamente. Este tremendo volumen de texto, en su mayoría no estructurado, no puede ser simplemente procesado y percibido por las computadoras. Por lo tanto, se requieren técnicas y algoritmos eficientes y efectivos para descubrir patrones útiles. La minería de texto es la tarea de extraer información significativa del texto, esta especialidad ha ganado atenciones significativas en los últimos años (Allahyari, y otros, 2017)

Los datos de texto son un buen ejemplo de información no estructurada, que es una de las formas más simples de datos que se pueden generar en la mayoría de los escenarios. El texto no estructurado es procesado y percibido fácilmente por los humanos, pero es mucho más difícil de entender para las máquinas (Feldman & Sanger, 2007).

Los servicios públicos, por ejemplo, generan grandes cantidades de datos que pueden ser almacenadas en bases de datos estructuradas, tal es el caso del Servicio del 911 de la provincia de Jujuy. En este servicio, la Base de Datos principal cuenta con más de 5.000.000 de registros. Cada vez que se realiza un llamado a este servicio, es grabado, catalogado y registrado por un operario. El mismo le asigna un tipo de incidente y un color según la severidad del mismo, el color varía entre celeste, amarillo, verde y rojo, yendo de severidad baja a alta según corresponda.

La tabla principal donde se almacenan los llamados, posee una gran cantidad de campos donde el operario puede identificar el origen de la llamada, persona que llama, dirección, tipo de incidente, severidad, la dirección para hacer posible la georeferenciación entre muchos otros.

Un campo llamado *descripción* se emplea para dotar de mayor detalle la llamada recibida, al ser un campo libre y extenso, el operario se explaya con los detalles del llamado, por ejemplo puede introducir cuestiones como el origen o causa del incidente. Es decir, si una persona llama por violencia de género, el operario puede indicar que la causa del llamado se debe a que la persona que cometió la falta estaba en estado de ebriedad, en este caso el llamado queda asentado como "violencia de género", etiqueta de color "roja", pero no queda indicado en otros campos el origen del incidente, no queda registrado de forma paramétrica que la persona que inició el hecho violento se encontraba ebria, salvo el texto en el campo descripción.

Poder analizar dicho campo, empleando técnicas de minería de texto y de clasificación de información, permitirían obtener muchísimos detalles hasta ahora descuidados, además se ofrecería información que los reportes habituales y estadísticos que se obtienen desde la plataforma que emplean lo pasan por alto totalmente, al ser un campo adicional y al ser de "solo texto", la estadística descriptiva no es capaz de analizarlo.

La minería de texto ha sido ampliamente utilizada para hallar relaciones y patrones entre datos no estructurados (Satyabrata, Mangal, Jinse, Ki-Won, & Hee-Cheol, 2017; Zhong, Li, & Wu, 2012; Kwon, Kim, & Park, 2017), el empleo de técnicas de Procesamiento del Lenguaje Natural es una de las principales tecnologías que favorecen en gran medida el procesamiento de este tipo de información (Kao & Poteet, 2007; Thessen, y otros, 2018; Joshi, Macwan, Mistry, & Mahida, 2018), es por ello que los sistemas basados en este enfoque aportan un valor agregado a los datos que manejan. En ese mismo sentido, técnicas de inteligencia artificial como aprendizaje automático y aprendizaje profundo, han sido utilizadas en combinación a la minería de datos y de textos

para procesar información relacionada con la seguridad y cyber-seguridad (Dua & Du, 2016).

Clasificar los llamados atendiendo a este campo, rico en información, dará acceso a un conjunto de palabras claves, indicadores y datos que complementarán los informes periódicos, esta "nueva información" (que no es nueva porque siempre estuvo presente pero no en un formato accesible) abrirá la posibilidad de manejar informes detallados y específicos, atendiendo no solo a la causa u origen del llamado origen, sino también a las relaciones entre los eventos/sucesos. Este tipo de relaciones tiene que ver con los patrones que se logren identificar (Wang, Rudin, Wagner, & Sevieri, 2015).

Las técnicas de *text mining* implican el uso de técnicas de inteligencia artificial como el PLN (MANNING & SCHÜTZE, 1999) y Aprendizaje Automático (*Machine Learning*) (Witten, Frank, Hall, & Pal, 2016). Estas técnicas permiten por un lado extraer datos relevantes y por el otro encontrar patrones y relaciones entre los datos. Para ello es necesario desarrollar un modelo de predicción, separar los conjuntos de datos en datos de entrenamiento, prueba y validación para asegurar que las clasificaciones y agrupaciones sean correctas.

Una vez que los datos sean procesados, los modelos entrenados y validados, se empelarán técnicas de inteligencia artificial para visualizar las relaciones encontradas. Puesto que se parte de la hipótesis de que estos datos sin analizar, los datos contenidos en el campo descripción, aportarán nueva información a los reportes e informes que se utilicen. Como se mencionó anteriormente, esta nueva información no reemplazará el manejo actual sino que lo complementará y potenciará.

Interfaces basadas en texto requieren esfuerzo cognitivo para entender su contenido informativo. La Visualización de Información (InfoVis) tiene por objeto presentar la información visualmente, en esencia, para reducir la carga de trabajo cognitivo al sistema perceptivo visual humano (Ware, 2004; Ware, 2008).

La Visualización de Información abarca las técnicas de visualización que tienen que ver principalmente con datos abstractos, es decir, los datos para los cuales el usuario no tiene un modelo mental preconcebido. Por esta razón, la interacción es especialmente importante en InfoVis, ya sea para la exploración, análisis y/o presentación de los datos (Kosara, Hauser, & Gresh, 2003). La interacción permite al usuario implícitamente formar modelos mentales de las correlaciones y las relaciones entre los datos, a través del reconocimiento de patrones.

La utilización de representaciones estáticas como imágenes o gráficos sin interacción han quedado atrás, el usuario de hoy en día necesita interactuar con la representación presentada, la herramienta de soporte debe brindar las facilidades para intercambiar formas, colores y modos de representar la misma información, con lo cual la visualización de información en este proyecto no debe ser descuidada puesto que aportará una visión complementaría al análisis que se pretende realizar.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación propone estudiar, evaluar y aplicar técnicas de Minería de Textos para analizar la información no estructurada de la Base de Datos del Servicio del 911 de la provincia de Jujuy. Para ello se pretende emplear técnicas de Procesamiento del Lenguaje Natural para hallar relaciones ocultas entre los datos y no visibles con simples informes o consultas SQL. Las relaciones permitirán conocer y formar una nueva imagen a partir de los datos existentes, identificando en muchos casos el origen o la causa de un llamado y no solo la consecuencia del suceso.

3. RESULTADOS OBTENIDOS/ESPERADOS

El presente proyecto tiene como objetivo general analizar, profundizar y rastrear información no estructurada almacenada en la base de datos del Servicio del 911 de la provincia de Jujuy, en busca de datos que aporten mayor detalle y definición a los reportes e informes estadísticos que el sistema actual genera.

Particularmente se espera lograr:

- Analizar la base de datos actual en busca de información no estructurada y que no es tenida en cuenta en los reportes que el sistema entrega.
- Extraer características de datos no estructurados utilizando técnicas de Procesamiento del Lenguaje Natural (PLN).
- Diseñar, testear y calibrar esquemas de PLN en base a las características encontradas.
- Diseñar modelos de clasificación y agrupación de información mediante la aplicación de técnicas de Aprendizaje Automático.
- Testear, calibrar y comparar los distintos modelos de clasificación y agrupación diseñados.
- Representar las distintas relaciones y/o patrones entre los datos procesados mediante el empleo de técnicas de visualización de información.

4. FORMACIÓN DE RECURSOS HUMANOS

El presente proyecto está a cargo del Dr. José Federico Medrano como director, además cuenta con el apoyo de los docentes investigadores: Ing. César Castillo, Ing. Mario Tejerina y el Analista Programador Universitario Juan Carlos Rodríguez.

Este proyecto brinda un marco para que docentes auxiliares y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico.

Actualmente, se están dirigiendo dos trabajos finales de la carrera Ingeniería Informática (UNJu y UCSE DASS) relacionadas con la temática propuesta, tres alumnos pasantes colaboran y llevan a cabo tareas a lo largo de todo el proyecto y además participan en un proyecto de vinculación de la SPU, con el objetivo de iniciarse en la investigación.

5. BIBLIOGRAFÍA

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint*(arXiv:1707.02919).

Dua, S., & Du, X. (2016). *Data mining and machine learning in cybersecurity.* Auerbach Publications.

Feldman, R., & Sanger, J. (2007). *The text mining handbook: advanced approaches in analyzing unstructured data.* Cambridge university press.

Joshi, B., Macwan, N., Mistry, T., & Mahida, D. (2018). Text Mining and Natural Language Processing in Web Data Mining. *2 International Conference on Current Research Trends in Engineering and Technology*, *4*, págs. 392-394.

Kao, A., & Poteet, S. R. (2007). *Natural Language Processing and Text Mining.* Springer Science & Business Media.

Kwon, H., Kim, J., & Park, Y. (2017). Applying LSA text mining technique in envisioning social impacts of emerging

technologies: The case of drone technology. *Technovation, 60*, 15-28.

MANNING, C., & SCHÜTZE, H. (1999). *Foundations of statistical natural language processing.* MIT Press.

Satyabrata, A., Mangal, S., Jinse, P., Ki-Won, C., & Hee-Cheol, K. (2017). A text mining approach to identify the relationship between gait-Parkinson's disease (PD) from PD based research articles. *Proceedings of the International Conference on Inventive Computing and Informatics.*

Thessen, A., Preciado, J., Jain, P., Martin, J., Palmer, M., & Bhat, R. (2018). Automated Trait Extraction using ClearEarth, a Natural Language Processing System for Text Mining in Natural Sciences. *Biodiversity Information Science and Standards*.

Wang, T., Rudin, C., Wagner, D., & Sevieri, R. (2015). Finding patterns with a rotten core: Data mining for crime series with cores. *Big Data, 3*(1), 3-21.

Ware, C. (2004). *Information Visualization - Perception for Design.* Morgan-Kaufmann.

Ware, C. (2008). *Visual Thinking for Design.* Morgan Kaufman/Elsevier.

White, H., & McCain, K. W. (1989). Bibliometrics. *Annual Review of Information Science and Technology, 24*, 119-186.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques.* Morgan Kaufmann.

Zhong, N., Li, Y., & Wu, S. T. (2012). Effective pattern discovery for text mining. *IEEE transactions on knowledge and data engineering, 24*(1), 30-44.

Computación

Gráfica, Imágenes y

Visualización

ANÁLISIS VISUAL DE DATOS ESPACIO-TEMPORALES

M. Luján Ganuza[1,2], Leandro Luque[1], Osvaldo E. Agamennoni[2] y Silvia M. Castro[1,2]

[1] Laboratorio de I+D en Visualización y Computación Gráfica (VyGLab)
(UNS-CIC Prov. de Buenos Aires)
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
mlg@cs.uns.edu.ar, leandroluque.tw@gmail.com, smc@cs.uns.edu.ar

[2] Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC)
Departamento de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur (DIEC-UNS)
oagamen@uns.edu.ar

RESUMEN

La disciplina de la visualización abarca el desarrollo de métodos y herramientas visuales a los efectos de facilitar el manejo de grandes volúmenes de datos de manera rápida y eficiente.

Debido al amplio campo de aplicación de los datos espacio-temporales y a las variadas hipótesis y necesidades de investigación que se pueden plantear, se desarrollaron diferentes alternativas de análisis de este tipo de datos tales como algoritmos estadísticos, algoritmos de edición de cadenas, técnicas relacionadas con visualización, etc. En tanto los métodos estadísticos proveen resultados cuantitativos, las técnicas de visualización permiten que los investigadores analicen exploren diferentes niveles y aspectos de sus datos. El objetivo de esta Línea de Investigación es el diseño y desarrollo de técnicas de Visualización que permitan analizar tanto los aspectos espacio-temporales de los datos provenientes de eye-trackers como las complejas relaciones que puedan existir entre éstos, durante el proceso de lectura.

Palabras Clave: Datos Espacio-Temporales, Análisis Visual de Datos, Eye Trackers

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación (DCIC), de la Universidad Nacional del Sur (UNS) y en estrecha colaboración con el Laboratorio de Desarrollo en Neurociencias Cognitivas (LDNC, UNS-CIC). Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La visualización es una herramienta muy valiosa que nos permite detectar, inferir y sacar conclusiones sobre las posibles relaciones existentes entre los datos [1].

Los datos espacio-temporales son aquellos datos en los que los cambios a lo largo del tiempo o los aspectos temporales juegan un rol central o son de interés y además tienen

asociada información relacionada con su ubicación en cierto dominio espacial. En lo que respecta a los atributos espaciales, la estrategia básica para su visualización es potencialmente directa ya que éstos se mapean directamente en ese dominio espacial ya definido y se presentan en pantalla. En relación a sus atributos temporales, la visualización no es directa. Para crear representaciones visuales eficientes y que permitan el razonamiento sobre el tiempo, es preciso que los métodos de visualización también tengan en cuenta las características particulares de tiempo, esto es, su estructura jerárquica y potencialmente periódica. En este contexto, para datos que además de ser temporales, son espaciales, se reafirman los desafíos presentados para datos únicamente temporales, ya que en estos últimos, el mapeo espacial está implícitamente especificado.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea de investigación en particular, nos enfocaremos en un tipo específico de datos espacio-temporales, los datos provenientes de un registrador de movimientos oculares eye-tracker (ET). El ET registra y graba, durante un determinado tiempo y a una dada frecuencia tanto la posición ocular como otras variables adicionales (velocidad, aceleración y diámetro de la pupila, entre otras). El seguimiento de los movimientos oculares (eye-tracking) permite analizar la información adquirida por una persona durante la realización de diversas actividades tales como la lectura, la observación de una imagen, la conducción de un vehículo, etc.

Debido a los diversos campos de aplicación en los que se utiliza el seguimiento ocular, han sido desarrollados diferentes enfoques para analizar los datos provenientes de éste [2, 3].

En tanto los métodos estadísticos para el análisis de este tipo de datos proveen resultados cuantitativos [4, 5], las técnicas de visualización permiten que los investigadores analicen y exploren diferentes niveles y aspectos de los datos generados en sus experimentos llevados a cabo con el ET. Las técnicas de Visualización ayudan a analizar tanto los aspectos espacio-temporales de los datos como las relaciones complejas que puedan existir entre los diversos datos y tipos de datos obtenidos [6, 7, 8, 9]. Debido a la creciente complejidad de las tareas y estímulos posibles en los experimentos de eye-tracking, consideramos que la visualización jugará un rol cada vez más relevante en el análisis de este tipo de experimentos.

La lectura constituye una actividad muy propicia para estudiar las relaciones existentes entre los movimientos oculares y los procesos cognitivos vinculados con la memoria. A los efectos de introducir información al cerebro a través del sistema visual, los ojos se mueven rápidamente, llevando la información a una pequeña depresión de la retina (fóvea) donde la agudeza visual es mayor. En personas sanas, la duración de estos movimientos sacádicos varía entre 50 y 150 milisegundos. Entre un sacádico y otro se produce la fijación ocular. Las fijaciones duran entre 100 y 700 milisegundos, y es durante las mismas que la información visual ingresa al cerebro. El tiempo utilizado para procesar la información adquirida en cada fijación será mayor o menor, dependiendo de cuán relevante sea la información adquirida en relación a la tarea que se está efectuando.

El análisis de los registros de los movimientos oculares durante el proceso de lectura permiten evaluar, entre otras cosas, la capacidad cognitiva y/o el deterioro cognitivo temprano de manera objetiva, no invasiva y a un bajo costo [10, 11, 12].

Considerando lo expuesto y a los efectos de analizar el comportamiento ocular de una persona ante distintas situaciones (lectura de distintos tipos de texto, observación de una secuencia de imágenes, conducción de un vehículo, etc.), se han desarrollado distintos tests estadísticos sobre los datos obtenidos por investigadores. En este punto, y dada la gran interrelación de las variables analizadas se considera que sería de gran utilidad evaluar la existencia de patrones y entender las posibles relaciones existentes en la gran cantidad de información obtenida a partir de los movimientos oculares y las escenas sobre las cuales se generaron. Para lograr este objetivo se estudiarán, evaluarán y desarrollarán distintas alternativas de visualización de la información obtenida, tanto para información con características espacio-temporales obtenida directamente de los experimentos como para la información obtenida a partir de los procesos de clasificación y estadísticos de la información. Por otro lado es necesario diseñar y desarrollar métodos y herramientas de visualización que permitan analizar la extensa y compleja información obtenida.

3. RESULTADOS OBTENIDOS Y ESPERADOS

En este contexto, se espera que esta línea de investigación contribuya al desarrollo de tecnologías y soluciones en torno al análisis visual de datos espacio-temporales. Particularmente, nos centraremos en la visualización de datos provenientes de ETs, de características espacio-temporales, con el objetivo de evaluar el comportamiento de diversas personas durante el proceso de lectura. Este objetivo se alcanzará mediante el diseño y desarrollo de técnicas y herramientas para la visualización de datos provenientes de ETs durante el proceso de lectura de oraciones y otro tipo de textos.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis en desarrollo relacionadas con las líneas de investigación presentadas, así también como los proyectos de investigación y los becarios/as vinculados/as.

Tesis en Desarrollo: “Desarrollo de modelos del comportamiento ocular”, alumno: Juan Andrés Biondi. Tesis de Doctorado en Ciencias de la Computación. Directora: Dra. Silvia Castro, Dr. Osvaldo Agamennoni.

Becario: Leandro Luque. Denominación del plan de trabajo: “Análisis Visual de Datos provenientes de Registradores de Movimientos Oculares”. Beca doctoral CONICET 2018, adjudicada a partir de abril de 2019.

Ingreso a Carrera de Investigadora Científica: M. Luján Ganuza. Denominación del plan de trabajo: “Análisis Visual de Datos Espacio-Temporales”. Incorporación a la Carrera del Investigador Científico y Tecnológico de la Comisión de Investigaciones Científicas de la provincia de Buenos Aires (CIC) en la convocatoria 2018. Cargo en proceso de efectivización.

Proyecto: PGI 24//N037 “Análisis Visual de Grandes Conjuntos de Datos”. Directora: Silvia M. Castro.

Proyecto: PICT 2016-1009. “Modelado de la Dinámica Ocular para la Evaluación del Desempeño Cognitivo” (ANPCyT, 2016-2019) de la categoría Plan Argentina Innovadora 2020, dirigido por el Dr. Osvaldo Agamennoni.

5. BIBLIOGRAFÍA

[1] Cleveland, W. S. (1993). Visualizing data. Hobart Press.

[2] Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. OUP Oxford.

[3] Andrienko, G., Andrienko, N., Burch, M., & Weiskopf, D. (2012). Visual analytics methodology for eye movement studies. IEEE Transactions on Visualization and Computer Graphics, 18(12), 2889-2898.

[4] Dorr, M., Martinetz, T., Gegenfurtner, K. R., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. Journal of vision, 10(10), 28-28.

[5] Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. Behavior research methods, 45(3), 842-856.

[6] Kurzhals, K., Hlawatsch, M., Heimerl, F., Burch, M., Ertl, T., & Weiskopf, D. (2016). Gaze stripes: Image-based visualization of eye tracking data. IEEE TVCG, 22(1), 1005-1014.

[7] Kurzhals, K., Heimerl, F., & Weiskopf, D. (2014, March). ISeeCube: Visual analysis of gaze data for video. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 43-50). ACM.

[8] Pfeiffer, T. (2012, March). Measuring and visualizing attention in space with 3D attention volumes. In Proceedings of the Symposium on Eye Tracking Research and Applications (pp. 29-36). ACM.

[9] Goggins, S. P., Schmidt, M., Guajardo, J., & Moore, J. (2010, January). Assessing multiple perspectives in three dimensional virtual worlds: eye tracking and all views qualitative analysis (AVQA). In System Sciences (HICSS), 2010 43rd Hawaii International Conference on (pp. 1-10). IEEE.

[10] Fernández, G., Laubrock, J., Mandolesi, P., Colombo, O., & Agamennoni, O. (2014). Registering eye movements during reading in Alzheimer's disease: difficulties in predicting upcoming words. Journal of clinical and experimental neuropsychology, 36(3), 302-316.

[11] Fernández, G., Manes, F., Rotstein, N. P., Colombo, O., Mandolesi, P., Politi, L. E., & Agamennoni, O. (2014). Lack of contextual-word predictability during reading in patients with mild Alzheimer disease. Neuropsychologia, 62, 143-151.

[12] Fernández, G., Shalom, D. E., Kliegl, R., & Sigman, M. (2014). Eye movements during reading proverbs and regular sentences: The incoming word predictability effect. Language, Cognition and Neuroscience, 29(3), 260-273.

[13] Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., ... & Snyder, P. J. (2013). The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. FOCUS, 11(1), 96-106.

ANÁLISIS VISUAL DE DATOS MULTIDIMENSIONALES PROVENIENTES DE LAS CIENCIAS GEOLÓGICAS

M. Luján Ganuza[1], Antonella S. Antonini[1], Florencia Gargiulo[2], Gabriela Ferracutti[2], Ernesto A. Bjerg[2], Silvia M. Castro[1], Krešimir Matković[3], Eduard Gröller[4]

(1) Laboratorio de I+D en Visualización y Computación Gráfica
Dpto. de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)
mlg@cs.uns.edu.ar; antoo.antonini@gmail.com; smc@cs.uns.edu.ar
(2) INGEOSUR, Dpto. de Geología, Universidad Nacional del Sur
florenciagargiulo@gmail.com; gferrac@uns.edu.ar; ebjerg@ingeosur-conicet.gob.ar
(3)VRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH,
Donau-City-Strasse 11, 1220 Viena, Austria.
Matkovic@vrvis.at
(4) Technische Universität Wien, Institut für Computergraphik und Algorithmen,
Favoritenstrasse 9-11 / E186, A-1040 Viena, Austria
groeller@cg.tuwien.ac.at

RESUMEN

Muchos de los problemas a los que debemos enfrentarnos en diversas áreas del conocimiento implican el tratamiento de grandes volúmenes de datos que requieren ser analizados desde múltiples dimensiones. Al analizar datos multidimensionales el problema radica en considerar una gran cantidad de variables y sus relaciones simultáneamente. A medida que la cantidad de variables aumenta, la habilidad del usuario de entender la interacción o correlación entre éstas se ve severamente limitada si se emplean técnicas de visualización estándar, dado que éstas dimensiones no solo pueden referirse a escalas diferentes y representar distintos tipos de datos, sino que además, la calidad de estos datos puede no ser la esperada.

En este contexto, el objetivo de la presente línea de investigación es analizar la adecuación y/o la generación de técnicas de visualización de datos multidimensionales, para responder a las necesidades de visualización, volumen de datos, dimensionalidad, complejidad y capacidad exploratoria presentes en los datos geológicos multidimensionales.

Palabras Clave: Datos Multidimensionales, Análisis Visual de Datos, Visualización de Datos Geológicos, Rendering de Volúmenes

CONTEXTO

Este trabajo se lleva a cabo con investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación (DCIC-UNS), el Departamento de Geología de la UNS, el INGEOSUR CCT-CONICET, el VRVis Research Center y la Universidad Técnica de Viena (TU-Wien), Austria.

Los trabajos realizados bajo esta línea involucran a docentes investigadores, becarios doctorales y alumnos de grado.

1. INTRODUCCIÓN

La visualización es una herramienta muy valiosa que nos permite detectar, inferir y sacar conclusiones sobre las posibles relaciones existentes entre los datos. Esta herramienta es muy utilizada en diversas áreas de las ciencias, la ingeniería y la medicina, entre otras, en las cuales se generan volúmenes de datos cada vez más grandes y difíciles de analizar y comprender sin un soporte visual adecuado [1, 2, 3, 4, 5]. Es en estos casos en los que el aporte de la visualización a la exploración y entendimiento de grandes conjuntos de datos resulta altamente significativo.

Al aumentar la dimensionalidad de los datos, lograr una visualización eficiente de los mismos se convierte en una tarea cada vez más compleja, ya que los espacios multidimensionales son difíciles de entender, por su propia naturaleza, debido a la gran cantidad de dimensiones que se deben representar en un espacio muy limitado. El problema radica en la necesidad de considerar una gran cantidad de variables y sus relaciones simultáneamente. Muchas de las técnicas tradicionales sólo producen una visión parcial de los datos, ya sea por su dimensionalidad y/o su cantidad. Esto dificulta la detección, si fuese necesario, de las posibles relaciones que existen entre todas las variables que intervienen en un proceso, generando vacíos de conocimiento por la imposibilidad de tratar el fenómeno en toda su dimensionalidad.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En esta línea de investigación en particular, nos enfocaremos en un tipo específico de datos multidimensionales, los datos geológicos.

Los trabajos desarrollados por los geólogos, tanto de campo como de laboratorio, implican grandes desafíos en lo que respecta al procesamiento y análisis de los datos que se obtienen en cada uno de estos contextos. Una gran cantidad de estos datos son indudablemente datos multidimensionales.

Los geólogos analizan grandes volúmenes de datos provenientes de análisis geoquímicos. Estos análisis arrojan una gran cantidad de atributos para cada análisis efectuado y éstos tienen asociados dominios espaciales característicos como los prismas o los tetraedros. Debido a las distintas características de los datos involucrados, no hay una única metáfora visual que pueda satisfacer todas las necesidades analíticas. Por esta razón, resulta de suma utilidad contar con un conjunto de metáforas visuales y de técnicas de visualización asociadas que permitan analizar estos datos. Estas representaciones deben poder comunicar efectivamente las características del espacio de información así como incentivar el descubrimiento. En este contexto, y en base a los trabajos realizados y a la experiencia obtenida hasta el momento en lo que respecta a Visualización de Datos Geológicos, se propone avanzar en el conocimiento sobre el tema de análisis visual de datos multidimensionales constituidos principal, aunque no exclusivamente, por datos geoquímicos, isotópicos, petrológicos, mineralógicos y estructurales.

Las técnicas tradicionales de visualización, tales como scatter plots, histogramas, etc., sólo producen una visión parcial de la información. Actualmente existen técnicas de visualización utilizadas en distintos contextos, que permiten analizar grandes conjuntos de datos multidimensionales y que han probado ser de gran utilidad [6, 7, 8].

En este contexto, se plantea la necesidad de contribuir en el desarrollo de tecnologías y soluciones en torno al análisis visual de datos multidimensionales. Particularmente, nos centraremos en la visualización de datos geológicos, con el objetivo de desarrollar nuevos modelos, técnicas y herramientas que permitan a profesionales de las Geociencias explorar efectiva y eficientemente sus espacios de datos. Si bien nos enfocaremos en la aplicación de estas nuevas tecnologías en el campo de la Geología, estos métodos redundarán muy probablemente en soluciones extensibles a distintas disciplinas en las que también deba obtenerse información a partir de conjuntos de datos multidimensionales.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Los autores de este artículo conforman un grupo interdisciplinario integrado por profesionales de Ciencias de la Computación y de Ciencias Geológicas, que ha trabajado en temas relacionados con la Visualización de Datos aplicada a Geociencias desde hace varios años [9, 10, 11]. Además, con la colaboración del Institute of Computer Graphics and Algorithms de la Universidad Técnica de Viena (TU-Wien) y del VRVis Research Center de Viena, se han desarrollado nuevas técnicas de visualización para datos geológicos aplicados a la mineralogía y a la prospección de recursos naturales [12, 13, 14].

Específicamente, se plantea el diseño y desarrollo de técnicas y herramientas para la visualización de datos geoquímicos que contribuyan efectivamente a lograr una mejor comprensión de la interacción entre los procesos geológicos y las composiciones de minerales y de rocas en la configuración geológica de una región en particular. Se propone utilizar todos los atributos disponibles de los datos y metadatos generados a partir de éstos, para validar y mejorar técnicas utilizadas en el contexto de datos altamente dimensionales (por ejemplo, datos de química mineral y datos geoquímicos de rocas). Para alcanzar este objetivo, nos centraremos en dos líneas principales:

- Diseño y desarrollo de técnicas de visualización para datos multidimensionales en el contexto de datos geológicos. Analizar técnicas y métodos de análisis visual para incorporar más de 8 miembros finales (y hasta un total de 22), que son los que se consideran actualmente y también datos geoquímicos de elementos mayoritarios. Esto se corresponde con el diseño y desarrollo de nuevas técnicas de visualización para el análisis de minerales del grupo de los espinelos.
- Extensión de las técnicas avanzadas de clasificación de espinelos. Extender al espacio 3D las técnicas avanzadas de clasificación de los minerales del grupo de los espinelos desarrolladas en nuestro trabajo previo [13, 14] y definir las interacciones asociadas. Esto también contribuirá significativamente al desarrollo de tecnologías y soluciones en torno a la generación de visualizaciones interactivas en 3D.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo concerniente a la formación de recursos humanos se detallan las tesis finalizadas y en desarrollo relacionadas con las líneas de investigación presentadas, así también como los becarios y los proyectos de investigación vinculados.

Tesis Finalizada: “Interacciones en Visualización”, tesis de doctorado en Ciencias

de la Computación. Alumna: M. Luján Ganuza. Directora: Dra. Silvia Castro.

Proyecto Final Finalizado: "Visualización de Datos Geológicos en la Web". Proyecto final de la Ingeniería en Sistemas de Computación. Alumna: Antonella S. Antonini. Directora Silvia Castro. Codirectora: M. Luján Ganuza.

Proyecto Final Finalizado: "Técnicas 3D para Visualización de Datos Geológicos en la Web". Proyecto final de la Ingeniería en Sistemas de Computación. Alumno: Gonzalo Picorel. Directora Silvia Castro. Codirectora: M. Luján Ganuza.

Becaria: Antonella S. Antonini. Título del plan propuesto: "Visualización de Datos Geológicos en la Web". Beca de entrenamiento 2018. Comisión de Investigaciones Científicas de la Provincia de Buenos Aires. Adjudicada a partir de abril de 2018 y por un término de 12 meses.

Proyecto: PGI 24//N037 "Análisis Visual de Grandes Conjuntos de Datos". Directora: Silvia M. Castro.

Proyecto: PICT 2017-1246 "Análisis Visual de Datos en Geociencias". Directora: Silvia M. Castro.

Becaria: M. Luján Ganuza. Título del plan propuesto: "Análisis Visual de Datos Multidimensionales en Espacios Ad-Hoc". Beca Posdoctoral 2018 CONICET. Comisión Nacional de Investigaciones Científicas y Técnicas de la República Argentina. Adjudicada a partir de abril de 2019 y por un término de 24 meses.

5. BIBLIOGRAFÍA

[1] Cleveland, W. S. Visualizing Data. Hobart Press, 1993.

[2] Card, S. K., Mackinlay, J. D., and Shneiderman, B. Readings in information visualization: using vision to think. Morgan Kaufmann, 1999.

[3] Ware, C. Information Visualization: Perception for Design. Morgan Kaufmann, 2004.

[4] Spence, R. Information Visualization: Design for Interaction. 2007.

[5] Ward, M. O., Grinstein, G., and Keim, D. Interactive data visualization: foundations, techniques, and applications. CRC Press, 2010.

[6] J Pastizzo, M., Erbacher, R., and Feldman, L. Multidimensional data visualization. 158–62.

[7] Everitt, B., and Hothorn, T. Looking at multivariate data: Visualisation. In An Introduction to Applied Multivariate Analysis with R. Springer, 2011, pp. 25–60.

[8] Etemadpour, R., Linsen, L., Paiva, J. G., Crick, C., and Forbes, A. G. Choosing visualization techniques for multidimensional data projection tasks: A guideline with examples. In International Joint Conference on Computer Vision, Imaging and Computer Graphics (2015), Springer, pp. 166–186

[9] Ganuza, M. L., Castro, S. M., Martig, S. R., Ferracutti, G., and Bjerg, E. Mineral compositions visualization implementing the spinel prism. *Proceddings Congreso Argentino de Ciencias de la Computación. CACIC 2009* (2009), 576–585.

[10] Ganuza, M. L., Castro, S. M., Ferracutti, G., Bjerg, E. A., and Martig, S. Spinelviz: An interactive 3d application for visualizing spinel group minerals. Computers & Geosciences 48 (2012), 50–56.

[11] Ferracutti, G. R., Gargiulo, M. F., Ganuza, M. L., Bjerg, E. A., and Castro, S. M. Determination of the spinel group end-members based on electron microprobe analyses. Mineralogy and Petrology 109, 2 (2015), 153–160.

[12] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. The spinel explorer - interactive visual analysis of spinel group minerals. IEEE Trans. Vis. Comput. Graph. 20, 12 (2014), 1913–1922.

[13] Ganuza, M. L., Gargiulo, M. F., Ferracutti, G., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. Interactive semi-automatic categorization for spinel group minerals. In 2015 IEEE Conference on Visual Analytics Science and Technology, VAST 2015, Chicago, IL, USA, October 25-30, 2015 (2015), pp. 197–198.

[14] Ganuza, M. L., Ferracutti, G., Gargiulo, M. F., Castro, S. M., Bjerg, E. A., Gröller, E., and Matkovic, K. Interactive visual categorization of spinel-group minerals. Proceedings of the Spring Conference on Computer Graphics (2017)

Aplicaciones de la Televisión Interactiva y tecnologías afines para el mejoramiento de los pueblos latinoamericanos

María José Abásolo[12], Armando De Giusti[1], Marcelo Naiouf[1], Patricia Pesado[1], Magdalena Rosado[13], Joaquín Pina[4], Raoni Kulesza[5], Telmo Silva[6]

[1]Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, degiusti}@lidi.info.unlp.edu.ar

[2]Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

[3] Universidad Católica de Santiago de Guayaquil, Ecuador

[4] CUJAE, Cuba
jpina@ceis.cujae.edu.cu

[5] Universidad Federal de Paraiba, Joao Pessoa, Brasil
raoni@lavid.ufpb.br

[6] Universidad de Aveiro, Aveiro, Portugal
tsilva@ua.pt

Resumen

Desde el año 2012 diversas universidades iberoamericanas conforman la Red de Aplicaciones y Usabilidad de la Televisión Digital Interactiva (RedAUTI) coordinada por la UNLP. La línea de investigación y desarrollo presentada consiste en estudiar, desarrollar y evaluar aplicaciones de Televisión Digital Interactiva (TVDi) y tecnologías complementarias para el mejoramiento de los pueblos latinoamericanos. Uno de los principales objetivos es la formación de recursos humanos y fortalecimiento de la investigación mediante el trabajo intergrupal entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Televisión digital, TVDi, Interactividad

Contexto

La RedAUTI[1] cuenta con la participación de 29 universidades de España, Portugal y 11 países latinoamericanos (Argentina, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, Guatemala, Perú, Uruguay, Venezuela). La RedAUTI fue creada y financiada por el Proyecto "512RT0461 Red de Aplicaciones y Usabilidad de la Televisión Digital Interactiva (RedAUTI)" del Programa Iberoamericano de Ciencia y Tecnología para el Desarrollo (CYTED) durante el período 2012-2015. En 2016 un grupo de universidades nacionales fueron financiadas por el programa Redes IX de la Secretaría de Políticas Universitarias del Ministerio de Educación Argentino. Recientemente fue aprobado por la Oficina de Naciones Unidas para la Cooperación Sur-Sur con el Fondo Fiduciario Pérez Guerrero (PGTF INT/19/K08) el proyecto internacional: "TDTi*ABC La Televisión Digital Terrestre aplicada al mejoramiento de los pueblos de los países en vías de desarrollo".

Introducción

La televisión digital abre la oportunidad de crear aplicaciones interactivas, de ahí la denominación de Televisión Digital Interactiva (TVDi). La interactividad

[1] http://redauti.net

permite un diálogo donde el participante tiene la capacidad de intervenir en los programas o servicios que recibe. Con la TVDI el consumidor puede pasar de ser un espectador pasivo a convertirse en un participante activo. 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. 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 [1][2] [3][4][5][6].

En Argentina los primeros abordajes en materia de interactividad llegaron desde la investigación académica en universidades nacionales y en experiencias pedagógicas y de transferencia tecnológica. En los casos desarrollados, la noción de interactividad se relaciona con la navegación, la presentación de contenidos en formatos multimedia (gráficos, sonoros, audiovisuales y transferencia de datos). Sin embargo, y si bien se han desarrollado algunas pruebas, falta desarrollar de manera concreta la articulación del *set-top-box* con la conectividad a internet, establecer un canal de retorno y generar una interactividad plena donde el televidente pueda intervenir en la producción de los contenidos y sus intervenciones sean consideradas, tal como ha sucedido en otros países del mundo. A las experiencias que se desarrollan actualmente en nuestro país se las denomina de interactividad local, donde el usuario puede acceder a contenidos interactivos pero no puede enviar datos de vuelta. Los ejemplos más comunes de aplicaciones interactivas locales son las guías electrónicas de programación, el teletexto digital, o la información adicional.

Artola et al [7] presentan una recopilación de aplicaciones de TVDI que se están desarrollando en Argentina y se las clasificó de acuerdo a una serie de criterios específicos tales como 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. Esto da cuenta de la intervención de la universidades argentinas en esta temática.

El término t-learning se refiere al aprendizaje interactivo a través de un televisor, es decir aplicaciones interactivas para TVDI enfocadas principalmente a apoyar los procesos de enseñanza y aprendizaje. El uso de la TVDI en educación se encuentra aún en pleno desarrollo e investigación, en particular en los países latinoamericanos. Carrascal et al [8] presenta una revisión de proyectos y propuestas en torno a la Televisión Digital Interactiva (TVDI) y su aplicación en diversos escenarios de aprendizaje.

La TVDi 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 o no están acostumbrados a la tecnología. 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 [9]. 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 un potente mecanismo de acceso a la información que de otro modo no sería posible. Como la televisión está presente en la mayoría de los hogares y los adultos mayores interactúan con la misma, existe el reto de utilizar este dispositivo de interacción en el ámbito de la salud y de la asistencia, que permita ayudar al adulto durante su proceso de envejecimiento.

Líneas de investigación y desarrollo

La RedAUTI se propone el intercambio de aplicaciones interactivas desarrolladas para su

evaluación en los diferentes contextos. El objetivo general es apoyar la formación de los recursos humanos, estimular la investigación de forma transdisciplinaria y fortalecer los grupos de investigación de las Universidades iberoamericanas en temas relacionados con TVDI, avanzando en el diseño, implementación y despliegue de aplicaciones, servicios y producción de contenidos para TVDI, en sus múltiples plataformas, de código abierto, de interés colectivo, para dar solución a problemas del contexto iberoamericano.

El proyecto que se lleva a cabo tiene como objetivo demostrar las potencialidades de la TDT en los países Argentina, Brasil y Cuba mediante contenido multimedia y softwares adaptados a los escenarios priorizados para el desarrollo en las esferas: Educación, Sanidad, Alimentación y Patrimonio. El proyecto es un marco propicio para aplicar otras tecnologías emergentes como Realidad Virtual (RV) /Realidad Aumentada (RA) y dispositivos móviles para enriquecer las funcionalidades y contenidos presentados al televidente. También se tienen en cuenta modernas técnicas de minería de datos y su integración con redes sociales y portales corporativos para llegar al ciudadano con contenido útil, dinámico y personalizado.

Las líneas de investigación y desarrollo principales son:

- Contenidos y Aplicaciones Interactivas para TVDi, con especial interés en T-learning y Asistencia a Adultos Mayores
- Herramientas informáticas para el desarrollo de aplicaciones para TVDi
- Experiencias de usuario con la TVDi: usabilidad y accesibilidad
- Combinación de TVDi con otras tecnologías como RA/RV

Resultados y Objetivos

Los objetivos específicos de la RedAUTI son:

1- Identificar problemas del contexto iberoamericano que pueden ser solucionados con la construcción de contenidos, aplicativos y servicios soportados en las tecnologías de la TVDI. En particular, ofrecer soluciones de código abierto, de interés colectivo, asociadas a la mejora del bienestar de la población (t-educación, t-salud, t-gobierno, etc.) con énfasis en la inclusión social y digital.

2- Crear el "Foro Iberoamericano de TVDI" para incentivar el debate y la investigación conjunta.

3- Realizar un evento anual de difusión abierto para intercambiar resultados de la investigación en el desarrollo de aplicaciones interactivas y contenidos para TVDI.

4- Ofrecer cursos por parte de los investigadores integrantes de la RedAUTI en temas relacionados con el desarrollo de aplicaciones interactivas y contenidos para TVDI.

5- Incentivar las estancias de trabajo de doctorandos e investigadores en las instituciones participantes de países que se encuentren en estadio más adelantado de aplicaciones y desarrollo de contenidos.

7- Avanzar en la investigación transdisciplinaria para el desarrollo de aplicaciones, servicios y producción de contenidos sobre TVDI, creando grupos de trabajo y asociaciones estratégicas de investigación para buscar la creación de nuevos servicios en sectores emergentes como son por ejemplo los juegos y la educación, la TV social, inteligencia ambiental, etc.

Entre los principales resultados de la RedAUTI se resaltan:

- Difusión mediante el portal web y redes sociales con participación de los miembros enviando información para difusión como publicaciones, proyectos, noticias, eventos, etc.
- Organización de un evento científico anual denominado Jornadas de difusión y capacitación de Aplicaciones y Usabilidad de la TVDi desde el año 2012. Actualmente se está organizando el próximo evento jAUTI 2019 a realizarse en Río de Janeiro en el mes de noviembre de 2019 en concordancia con el evento WebMedia 2019.
- Publicación conjunta de libros CCIS Springer con la selección de los mejores artículos extendidos para ampliar el alcance de la difusión de los resultados de las investigaciones conjuntas [10][11][12][13].
- Se están desarrollando contenidos y aplicaciones de TVDi para asistir el

mejoramiento de las alteraciones de la marcha en personas mayores, con el objetivo de probar los desarrollos realizados en Ecuador

- Se planifica una estancia de investigación de dos profesores extranjeros en la UNLP para el mes de junio 2019

Formación de recursos humanos

En la actualidad se está desarrollando una tesis de doctorado en el marco de esta línea de investigación dirigida conjuntamente entre docente argentino y docente extranjero:
- 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)

Referencias

[1] Ferraz de Abreu, Jorge; Abásolo Guerrero, María José; Almeida, Pedro; Silva, Telmo (Eds.) (2017) Anales jAUTI'17. *6ª Jornadas de Aplicaciones y Usabilidade de la Televisión Digital Interativa.* UA Editora, ISBN 978-972-789-521-2

[2] Abásolo Guerrero, María José; Pina Amargós, Joaquín (2017) *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.* UNLP, ISBN: 978-950-34-1514-6

[3] A.Bibiloni, F. Perales, M.J.Abásolo (Eds.) (2016) *"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. Universitat de les Illes Balears (España). ISBN: 978-84-608-5567-5

[4] Abásolo, M.J., Kulesza, R. (Eds.) (2015) *"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

[5] Abásolo, M.J., Castro, C. de (Eds.) (2014) *"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

[6] Abásolo, M.J., De Giusti A. (Eds.) (2013) "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

[7] 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

[8] Carrascal, M.; Sanz, C.; Abásolo, M.J. (2017) *Revisión de experiencias de t-learning.* En Abásolo, M.; Pina Amargós, J. 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. UNLP, ISBN: 978-950-34-1514-6

[9] 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%A7a_Os%20tele-cuidados%20e%20a%20televis%C3%A3o%20interactiva_done.pdf

[10] Abásolo, María José; Ferraz de Abreu, Jorge; Almeida, Pedro; Silva, Telmo (eds.)

(2017) Communications in Computer and Information Science 813, Springer ISBN 978-3-319-90170-1 https://www.springer.com/us/book/9783319901695

[11] 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

[12] 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

[13] 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/9783319226552

Aplicaciones de Realidad Virtual, Realidad Aumentada e Interfaces Multimodales

María José Abásolo[1 2], Armando De Giusti[1], Marcelo Naiouf[1], Patricia Pesado[1],Cecilia Sanz[1 2], Sebastián Barbieri[3], Ramiro Boza[1], Wilma Gavilanes[1 4], Alejandro Mitaritonna[1 5], Nahuel Prinscich[1], Mario Alberto Vincenzi[1] , Francisco Montero[6], Francisco Perales López[97]

[1] **Instituto de Investigación en Informática LIDI (III-LIDI)**
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, degiusti, mnaiouf, ppesado, csanz}@lidi.info.unlp.edu.ar

[2] **Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)**

[3] **Facultad de Cs.Exactas.Universidad Nacional del Centrode la Pcia.de Bs.As.(UNICEN)**

[4] **Universidad Técnica de Ambato, Ecuador**

[5] **Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)**

[6] **Universidad de Castilla-La Mancha, España**
francisco.msimarro@uclm.es

[7] **Universidad de las Islas Baleares, España**
paco.perales@uib.es

Resumen

La línea de investigación y desarrollo presentada en este artículo realiza actividades de análisis, desarrollo y evaluación de aplicaciones de Realidad Virtual, Realidad Aumentada, Interfaces Multimodales y Visión por computador. Uno de los principales objetivos es la formación de recursos humanos y el fortalecimiento de la investigación mediante el trabajo intergrupal entre diferentes instituciones nacionales y extranjeras.

Palabras Clave: Realidad Aumentada, Realidad Virtual, Interfaces multimodales, Visión por Computador

Contexto

La investigación relacionada con Visión por Computador y Realidad Aumentada forma parte 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.

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 [1][2][3][4][5]. En este artículo se presentan los principales avances alcanzados en esas diferentes líneas de investigación.

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[6].

En esta línea de investigación se han analizado diferentes proyectos de RA en el ámbito militar, y se ha trabajado sobre la implementación del framework RAIOM (Realidad Aumentada para la Identificación de Objetivos Militares) desarrollado en el CITEDEF (Instituto de Investigaciones

Científicas y Técnicas para la Defensa) [7] [8][9][10].

RAIOM, apoyándose en la RA, utiliza dispositivos móviles, visión por computador y sensores externos para el reconocimiento, detección, ubicación, identificación y suministro de información contextual.

Con RAIOM se han diseñado las componentes de software o módulos que forman parte del *middleware* del framework que lleva asociado Como parte de la arquitectura del sistema propuesto se implementó procesamiento distribuido. El proceso distribuido planteado está soportado por el diseño de una arquitectura Cliente / Servidor. Del lado del cliente de ejecutan los procesos menos intensivos tales como:

- Captura de video
- Uso de sensores (ubicación, tracking)
- Comunicación
- Mapeo
- Renderizado de imágenes

Mientras que del lado del servidor se implementan procesos que soportan cargas intensivas de cómputo.

En particular se utilizó un procesador mini board ODROID-XU3 como servidor y gafas Epson Moverio en el cliente.

Las siguientes características funcionales son incluidas en RAIOM:

- Radar de 360 grados: los objetivos tácticos y de amenaza se visualizan en el radar mediante el uso de símbolos militares.
- Posición del operador: el norte magnético sirve para guiar al operador. La posición relativa del operador se calcula de acuerdo con sus movimientos.
- Símbolos militares adaptables: los símbolos utilizados son adaptables para identificar las amenazas y los objetivos tácticos.
- Reconocimiento de puntos de referencia: la aplicación es capaz de reconocer edificios y otros tipos de infraestructura.
- Reconocimiento de gestos: los datos bajo demanda se pueden obtener mediante el reconocimiento de gestos.
- Mapas interactivos: el operador puede ver el terreno mediante el reconocimiento de gestos. Los mapas se visualizan superpuestos sobre la mano del operador.

El despliegue se evaluó para medir la efectividad, la eficiencia y la facilidad de uso para cumplir con los tres niveles del modelo de conciencia situacional (CS) propuesto por Endsley [17], es decir, la percepción, la compresión y la proyección.

La provisión de métodos válidos y confiables para evaluar la CS es esencial durante el diseño y la evaluación del sistema. Se requieren técnicas válidas y confiables. Se utilizaron diferentes técnicas para llevar a cabo las actividades de evaluación, como User Testing, Thinking Aloud y SAGAT / SART [18] [19]. Además, la experiencia del usuario (UX) se evaluó mediante el cuestionario UMUX [20].

Realidad Aumentada para brindar Información al Ciudadano

Algunos autores como M.Lens-Fitzgerald [21] 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 Dirty[1], EWG'S Healthy Living[2], GoodGuide Scanner[3]. El problema de

[1] Think Dirty: https://www.thinkdirtyapp.com

[2] EWG's Healthy Living: https://www.ewg.org/apps/

[3] GoodGuide Scanner: https://www.goodguide.com/about/mobile

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 presenta en inglés.
Por ello se desarrolló la aplicación móvil ConZoom que brinda 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 barras presente en el producto. La aplicación permite conocer no sólo información nutricional a través de gráficos, sino también información sobre cuán ecológico es su envase [22].

Realidad Aumentada en Educación

La RA es una tecnología innovadora y con importantes oportunidades en los procesos de enseñanza-aprendizaje. Muchas investigaciones recientes analizan las posibilidades educativas alcanzadas, destacando entre ellas la presentación de contenidos interactivos y las mejoras en el rendimiento del proceso de enseñanza y aprendizaje [23]. Los recursos utilizados de RA, como videos y modelos 3D, visualizados a través de dispositivos móviles, de forma contextualizada, pueden ayudar a los estudiantes a comprender contenidos abstractos a la vez que resulta más motivante que las clases tradicionales en el aula.
La incorporación de la RA en los procesos educativos en diferentes disciplinas, incluyendo medicina, arquitectura, urbanismo, matemáticas y geometría, idiomas, arte, historia, ciencias naturales, química, física y geografía. Se está realizando una recopilación y catalogación de aplicaciones en diferentes áreas que pueden ser de utilidad a docentes de los diferentes niveles educativos.
Los recursos desarrollados en RA deben basarse en planteos educativos precisos con el objetivo de ser efectivos, relacionando así las habilidades cognitivas del alumno
También se exige que la producción de los recursos educativos digitales sean eficientes para ser empleados en las actividades del docente, con posibilidad de reusabilidad, abiertos y de calidad.

Interfaces Multimodales

En la actualidad, coexisten un conjunto de paradigmas de interacción persona-ordenador, tales como interfaces multimodales, afectivas, ubicuas, tangibles, RA y RV así com realidad mixta, basadas en agentes conversacionales, etc. que van más allá del paradigma tradicional WIMP. El punto en común en todas ellas es la búsqueda de una interacción natural y multimodal, a través de los diversos canales de comunicación humana, embebida en el propio entorno físico del usuario mediante la sensorización y actuación sobre el mismo. Lo que se está produciendo actualmente es una convergencia de paradigmas en los que el foco se pone en la experiencia de usuario seleccionando los modos de interacción y visualización más adecuados en cada caso. Sin embargo, el avance de estos paradigmas emergentes se está frenando debido a la complejidad del prototipado de aplicaciones que los soporten. La creación de aplicaciones informáticas que adopten estos paradigmas de interacción supone la integración de elementos físicos, virtuales e incluso sociales (estado de ánimo de los usuarios, comportamientos grupales, etc.), introduciendo así una nueva dimensión de complejidad al desarrollo de aplicaciones informáticas de la entrante era post-WIMP y retardando su avance.
Los nuevos paradigmas de interacción tienen un ámbito de aplicación perfecto en los videojuegos. Los juegos pervasivos están demostrando una gran potencialidad tanto en su vertiente lúdica como educativa. Adicionalmente, son sus características de espacialidad (mezcla de la información digital con la física de forma natural), sociabilidad (fomentando las actividades grupales) y de personalización (adaptándose a las características de los usuarios o a su contexto), las que abren la posibilidad de su uso a colectivos de niños con algún tipo de trastorno o dificultad con otros modelos de

interacción más tradicional (a través de periféricos o en entornos reglados).

En lo que respecta a agentes sociales y niños con *déficit de habilidades sociales (TDA, TDAH, etc)*, es fundamental potenciar y mejorar la intervención en edades tempranas, algunas estrategias pueden basarse en la utilización de robots sociales. Los niños que interactúan con robots manifiestan mejoras en las habilidades sociales y comunicativas, aumentan el lenguaje espontáneo, conectan con el sujeto y el entorno, desarrollan comportamientos sociales hacia los robots y reducen la frecuencia de comportamientos repetitivos y estereotipados [24][25].

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
- Realidad Virtual y Aumentada aplicada a la educación
- Interfaces multimodales aplicadas a la valoración de niños con TEA y THDA

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).
- Tras 4 años de investigación y pruebas de concepto el framework RAIOM ha concluido las últimas fases de diseño y pruebas.
- En el ámbito de la asistencia al ciudadano se finalizó el primer prototipo de aplicación móvil ConZoom 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.

Formación de recursos humanos

Las siguientes diferentes tesis de grado y postgrado se desarrollan en el marco de esta línea de investigación:

- Alejandro Mitaritonna. "Realidad Aumentada para la Identificación de Objetivos Militares" . Directores: M.J.Abásolo, M. Larrea, F. Montero(tesis de doctorado presentada para evaluación en marzo de 2019)
- Nahuel Prinsich y Ramiro Boza. Aplicación móvil para el consumo consciente y responsable. Directores: Sanz, C. y Abásolo M.J. (tesis de grado aprobada en febrero de 2019)
- 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)
- Natalia Encina. "Evaluación de browsers de realidad aumentada para apoyar procesos de enseñanza -aprendizaje" Director: Abásolo, M.J. (tesis de especialización en curso).
- Mario Vincenzi. "La Realidad Aumentada en la educación. Vigencia, proyecciones y límites" Director: Abásolo, M.J. (tesis de especialización 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)
-Martíin Becerra "Aumentación de Sistemas SCADA en el Contexto de la Industria 4.0" Director: Jorge Ierache (UNLaM) , María José Abásolo (UNLP).

Referencias

[1] Abásolo, M.J.; Mitaritonna, A; Castañeda, S.; Sanz, C.; Boza, R., Prinsich, N.; Silva, T., Rosado, M.; Naiouf, M.; Pesado, P.; De Giusti, A. (2018) *Aplicaciones de visión por computador, realidad aumentada y TVDi.* XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). p. 389-395, RedUNCI; ISBN: 978-987-3619-27-4

[2] Abásolo, M.J.; Sanz, C.; Naiouf, M.; De Giusti, A.; Santos, G.; Castro, M.;Bouciguez, M.J. G. (2017) *Realidad Aumentada, Realidad Virtual e Interacción Tangible para la Educación.* XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017), pp. 1312-1316, RedUNCI, ISBN 978-987-42-5143-5

[3] Abásolo, M.; Mitaritonna, A.; Bouciguez, M.; Encina, N.; Vincenzi, 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, RedUNCI, ISBN 978-950-698-377-2

[4] Abásolo, M; Mitaritonna A.; Encina N.; Vincenzi 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), RedUNCI, ISBN 978-987-633-134-0

[5] 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.* WICC 2014 XVI Workshop De Investigadores en Ciencias de la Computación, RedUNCI, ISBN 978-950-34-1084-4

[6] 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

[7] Mitaritonna, A.; Abásolo, M. J. (2015) *Improving Situational Awareness in Military Operations using Augmented Reality*. Proceedings of WSCG 2015. ISBN Nº:978-80-86943-72-5, 2013.

[8] Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves (2014) *RAIOM. Introducción a la arquitectura del framework de Realidad Aumentada.* VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014)

[9] Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves (2014) *RAIOM. Introducción a los algoritmos de visión por computador.* VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014)

[10] Mitaritonna, A.; Abásolo, M. J.(2013) *Mejorando la conciencia situacional en operaciones militares utilizando la realidad aumentada* (2013) Proceedings of XVIII Congreso Argentino de Ciencias de la Computación. ISBN Nº: 978-987-23963-1-2 pp. 356-365.

[17] M. R. Endsley. Design and evaluation for situation awareness enhancement. Santa Mónica, CA, USA, 1988.

[18] M. R. Endsley. Measurement of Situation Awareness in Dynamic Systems. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1), pp. 65- 84, 1995.

[19] M. R. Endsley. Situation Awareness Global Assessment Technique (SAGAT). Aerospace and Electronics Conference – Proceedings of the IEEE 1988 National, 3, pp. 789-795, 1988.

[20] K. Finstad. The Usability Metric for User Experience. Interacting with Computers, pp. 323-327, 2010.

[21] 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]

[22] Boza, R.; Prinsich, N., Abásolo, M.J.; Sanz, C. (2019) Aplicación móvil para el consumo consciente y responsable. Directores. Tesis de grado de la carrera Licenciatura en Informática, Facultad de Informática, Universidad Nacional de La Plata, febrero 2019.

[23] Gavilanes, W., Abásolo, M. J., Cuji, B. (2018) *"Realidad Aumentada en la Educación: una Revisión desde la Perspectiva Pedagógica",* Revista ESPACIOS. ISSN 0798 1015

[24]Alexander, E., Dickstein-Fischer, L., Fischer, G. S., Harrington, K., Su, H. & Yan, X. (2011). *An Affordable Compact Humanoid Robot for autism Spectrum Disorder Interventions in Children.* Engineering in Medicine and Biology Society (ed.), 33rd Annual International conference of the IEEE EMBS (p. 5319-5322). doi: 10.1109/IEMBS.2011.6091316

[25]Kim E., Lauren D., Berkovits E., Bernier D., Leyzberg, Shic, Rhea P., Scasselati B. (2012). *Social Robots as Embedded Reinforcers of Social Behavior in Children with Autism.* Journal Autism Dev Disord.

INTERFAZ HUMANO-MÁQUINA (DISEÑO UX)

Di Cicco Carlos[1], Jaszczyszyn Adrian[1], Luengo Pablo[1], Cicerchia Benjamín[2], Alvarez Eduardo[1], Charne Javier[1], Goitea Ezequiel[3], Blanco Matías[3]

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 N° 1119 3er Piso, Junín (B) – TE: (0236) 4477050 INT 11610

{carlos.dicicco, adrian.jaszczyszyn, pablo.luengo,benjamin.cicerchia, eduardo.alvarez, javier.charne, ezequiel.goitea, matias.blanco}@itt.unnoba.edu.ar

Resumen

El desarrollo de la interfaz apropiada en un producto de software o hardware es clave para que este producto cumpla con sus requisitos.
Durante la ejecución de proyectos tecnológicos desarrollados en el ITT, se ha comprobado que resulta difícil satisfacer los requerimientos no funcionales para la comunicación entre el humano y la máquina (*Human Machine Interface* - HMI), haciendo necesaria la producción de múltiples prototipos de un mismo producto. Los productos analizados en este documento tienen como propósito identificar, contextualizar, evaluar y desarrollar conceptos que permitan mejorar la toma de decisiones vinculadas a las HMI. Se describirán los distintos entornos en los que se está trabajando, entre ellos, el sistema de voto electrónico, GPS en el deporte, infrarrojo y captura de imágenes.

Palabras clave: Interfaces, Analfabetismo Digital, Sociología, Semántica, Semiótica, Software, Touchscreen, Ergonomía, Hardware, ATM.

Contexto

Esta línea de investigación "Interfaces Humano - Máquina" se enmarca en los proyectos "Informática y tecnologías emergentes" y "Tecnología y Aplicaciones de Sistemas de Software: Innovación en procesos, productos y servicios" que financia la UNNOBA y tiene como lugar de trabajo en el ITT.

1- Introducción

En la actualidad las computadoras son utilizadas por las personas para todo tipo de aplicaciones y objetivos de la vida cotidiana y empresarial. Los sistemas informáticos pueden llegar a tener poca intervención humana o tener la necesidad de ser accesibles permitiendo el uso de todo tipo de personas. [1, 2, 3]
Inmersos en esta realidad es donde se hace necesario avanzar en el desarrollo tecnológico y en la relación de uso y operatividad entre humano y máquina. [4, 5, 6]
De acuerdo Enrique Ruiz-Velasco Sánchez hay numerosas disciplinas que intervienen en las interacciones del hombre-máquina, donde

---

[1] Docente Investigador - ITT
[2] Becario Doctoral CIC - ITT
[3] Alumno - ITT
[4] ITT - Centro Asociado CIC

las ciencias cognitivas han evolucionado hacia una mejor comprensión y entendimiento mentales junto con la psicología, las neurociencias, la inteligencia artificial, la antropología, la filosofía, la lingüística, con un esfuerzo conjunto tratan de describir, explicar y predecir el qué y el cómo entienden los usuarios de un sistema computacional a una interfaz. [7]

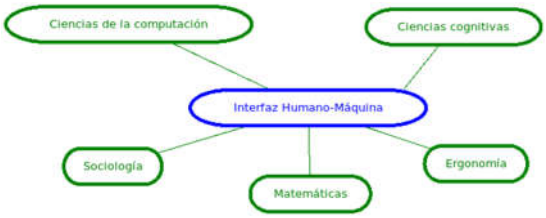


Fig.1: Disciplinas que intervienen en Interfaz Humano-Máquina

Analizando los sensores con los que se trabaja en cada proyecto (cardíaco, acelerómetro, GPS, infrarrojo, temperatura, humedad) los problemas más comunes y relevantes es la interpretación de la información recolectada al momento de ofrecer una interfaz que permita el análisis de la información de forma clara y concisa.

Se trabaja involucrando al usuario en el proceso de diseño y desarrollo, tratando de generar una conjunción entre la información que los distintos dispositivos ofrecen y lo que el usuario espera obtener y procesar para su posterior análisis. Además se busca obtener un diseño global e internacionalizado. [8]

Hoy en día hay equipos que cuentan con aplicaciones que ya incorporan diferentes tipos de sensores que miden magnitudes físicas, dando cuenta por ejemplo de los distintos movimientos realizados, su orientación, inclinación, etc. [9,10]

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.[11,12]

Es un desafío encontrar y obtener las interfaces adecuadas que permitan una simple lectura de la información proporcionada por los sensores, tales como rango, precisión, exactitud, sensibilidad, consumo de energía, Etc. datos necesarios para que las aplicaciones sean útiles para los distintos proyectos.[13,14]

Proyecto "GPS aplicado al deporte". El primer prototipo está formado por 5 módulos diferentes: microcontrolador, módulo GPS, sensor cardíaco, acelerómetro 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 encuentra en desarrollo una aplicación multiplataforma accesible y dinámica, que permitirá tomar decisiones en base a las distintas formas de visualización y análisis de los datos.

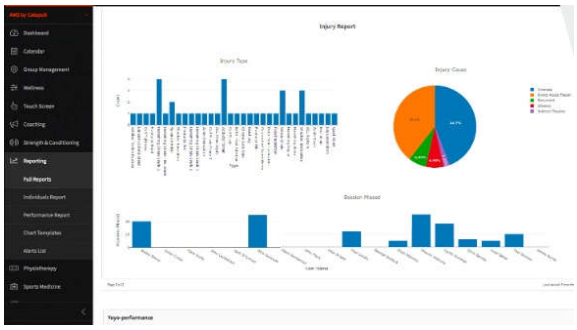


Fig.2: Interfaz GPS-Datos Estadísticos del Deportista

Proyecto "Sistema de Voto electrónico". Se trabaja en el armado de un nuevo prototipo de máquina de voto electrónico, con el objeto de ser más eficiente e interactiva con el votante.

Pretende mejorar la comunicación con la PC de administración de los presidentes de mesa, incorporando una pantalla táctil junto a la urna. El gabinete es portable permitiendo el traslado a distintas sedes de la Universidad. También se cambió la impresora, eliminando el mecanismo de arrastre de los comprobantes.

Fig. 3: Prototipo de la nueva máquina.

2- Líneas de Investigación y Desarrollo

La línea de investigación es definir las interfaces de software que permitan poder analizar y comprender 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. Comprensión del perfil de los usuarios, tareas y requerimientos futuros.

2. Diseño, definición del aspecto y estructura (objetos y acciones) de la interfaz.

3. Definición de estrategias de generación de prototipos, para poder realizar los test correspondientes.

4. Detección de problemas de usabilidad y accesibilidad. Documentación del prototipo, para su futura implementación.

5. Selección del modelo más eficiente para el problema planteado.

3- Resultados Obtenidos/Esperados

Se espera 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, cual es la mejor forma de mostrar la información generada, siempre centrado en el usuario. Para concluir en base a nuestra investigación y pruebas de campo cual consideramos más adecuada.
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 conformado por investigadores formados, investigadores en formación, 1 becarios de posgrado y 8 becarios de grado. Además se realizan convocatorias a alumnos avanzados de las

carreras del área de informática para que se inicien en tareas de investigación.

El desarrollo de esta línea de investigación espera contribuir a generar 2 Tesinas de Licenciatura en Sistemas, 1 Becario de intercambio internacional, 2 Práctica Profesional Supervisada.
Actualmente se encuentran en desarrollo 1 trabajo de Especialista, 2 Tesis de Magíster relacionadas con la línea de investigación.

5- Bibliografía

[1] Saltiveri, T.G. and Vidal, J.L. and Delgado, J.J.C., "Diseño de sistemas interactivos centrados en el usuario". Editorial UOC, S.L., (2011).

[2] Dix A., "Human-Computer Interaction", 3ª ed.. Pearson Education, (2008).

[3] Bach-Rita P. , Kercel S. W.," Sensory substitution and the human–machine interface", Trends in Cognitive Sciences Volume 7, Issue 12, Pages 541-546, (2003).

[4] Shneiderman, B. and Plaisant, C. and Cohen, M. and Jacobs, S., "Designing the User Interface: Strategies for Effective Human-Computer Interaction", 5ta edición, Pearson Education, (2017).

[5] Saariluoma, P. and Cañas, J.J. and Leikas, J., "Designing for Life: A Human Perspective on Technology Development",Palgrave Macmillan UK,(2016).

[6] Norman D.A. , Draper S.W., "User Centered System Design New Perspectives on Human-computer Interaction", CRC Press, (1986).

[7] Sánchez, E.R.V., "Educatrónica: innovación en el aprendizaje de las ciencias y la tecnología", Díaz de Santos, (2007).

[8] Alegría, A.F.S., Quiroz, J.D.C., "Evaluación colaborativa de la usabilidad en el desarrollo de sistemas software interactivos" , Universidad Autónoma de Occidente, (2016).

[9] Consejo Superior de Investigaciones Científicas (España), "Introducción a los sensores", C.S.I.C., (1987).

[10] Kymberly Tuck.,"Tilt sensing using linear accelerometers", Acelerometer Systems and Aplications Systems. Rev. 6, (2007).

[11] Pallás Areny, R, , "Sensores y Acondicionadores de señal", Barcelona-España, Boixareu, (2003).

[12] Fernández, C., García, A., "Sensores y actuadores" , Departamento de Publicaciones de la Escuela Universitaria de Informática de Madrid. (2009).

[13] Levin, M., "Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences Across Devices", O'Reilly Media, (2014).

[14] Bushe, Laura. "The Skeptic's Guide To Low-Fidelity Prototyping". Smashing Magazine. Recuperado de: http://www.smashingmagazine.com/2014/10/06/the-skeptics-guide-to-low-fidelity-prototyping/, (2014).

Investigación y desarrollo de aplicaciones de Realidad Aumentada y Visión por Computador utilizando un framework multipropósito

Alejandro Mitaritonna[1 2], Juan Lestani [1], Silvana Olmedo [1], Leonardo Inza [1]

[1] Laboratorio de Procesamiento de Imágenes y Visión Artificial
Universidad de Belgrano, Facultad de Ingeniería y Tecnología Informática
{alejandro.mitaritonna, juan.lestani, silvana.olmedo, leonardo.inza}@comunidad.ub.edu.ar

[2] Instituto de Investigaciones Científicas y Técnicas para la Defensa (CITEDEF)
amitaritonna@citedef.gob.ar

RESUMEN

En este artículo se presenta el proyecto de investigación 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. Dicho framework servirá como base para la construcción de aplicaciones móviles y de escritorio para ser utilizadas en diferentes disciplinas tales como medicina, entretenimiento, industria, turismo y educación, entre las principales. En el presente artículo se detallan los trabajos realizados hasta el momento, los logros alcanzados y las nuevas líneas de investigación que han surgido a medida que se desarrolla el proyecto.

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 ha desarrollado 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. Asimismo se está elaborando una tesis de grado de una de las integrantes del equipo de investigación.

INTRODUCCION

El proyecto está enmarcado en el área de procesamiento de imágenes y visión artificial [Mit17]. En [Mit18] se describe la línea de investigación presentada originalmente; las líneas/grupos de trabajo formados a fin de abordar las temáticas centrales: estimación de posición y orientación (pose) y registración de objetos, reconocimiento y clasificación de objetos y motores de renderizado de objetos; y los conceptos fundamentales de RA y VA.

El trabajo de la definición y construcción de un framework que facilite el desarrollo de

aplicaciones de RA comenzó por identificar y definir las partes componentes de un sistema de este tipo. De acuerdo a [Aba11] y basándose en lo que propone [Azu97], una aplicación de RA mínimamente está compuesto por las siguientes partes (figura 1):

- *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 móvil.

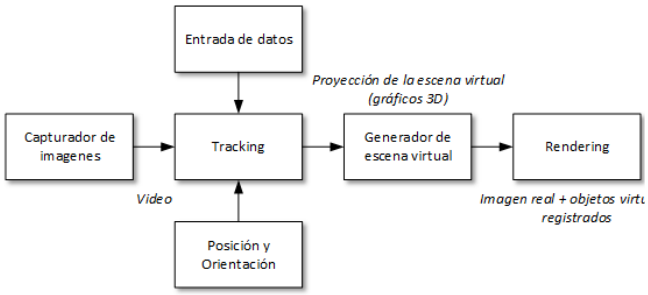


Figura 1

Adicionalmente se profundizó sobre los aspectos de VA, para ello se analizaron los componentes que forman parte del framework. Una de las particularidades del framework es que contará con un módulo de VA y cuya funcionalidad se centrará básicamente en el proceso de entender, interpretar y definir información visual de escenas en 2D y en 3D. Dicho módulo tendrá las siguientes características, entre las principales:

- Reconocimiento de Objetos
- Tracking de objetos basadas en características naturales
- Interfaz avanzada basada en gestos

Para concretar esta funcionalidad se está investigando y desarrollando pruebas de concepto cuyo objetivo es entender en profundidad la utilización de tecnologías de aprendizaje profundo para la clasificación de objetos. A continuación se detallan las principales líneas de investigación y desarrollo que el grupo de investigadores están realizando en el Laboratorio de Procesamiento de Imágenes y Visión Artificial de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano.

LINEAS DE INVESTIGACIÓN Y DESARROLLO

Modelo de referencia

Como primera aproximación al desarrollo de un framework se planteó que su arquitectura respondiera a las exigencias de un modelo de referencia para aplicaciones de RA [Mit04a]. Un modelo de referencia en ingeniería de sistemas y en ingeniería de software es un modelo que contiene un objetivo o idea básica de algo, y que se puede establecer como una referencia para múltiples propósitos. Un modelo de referencia consiste de un conjunto mínimo de conceptos, axiomas y relaciones propios de un dominio particular de problema, y es independiente de estándares específicos, tecnologías, implementaciones, o de cualquier otro detalle concreto. Un estudio de varios sistemas de RA reveló que a pesar de ser bastante diferentes en los detalles, la mayoría de los sistemas de RA existentes comparten una estructura arquitectónica en común. El modelo de referencia se organiza en seis subsistemas lógicos:

- *Subsistema de seguimiento (tracking)*: Responsable de responder a los cambios de posición y orientación del usuario.
- *Aplicación*: Responsable del control de flujo, la lógica de la aplicación y la coordinación.
- *Modelo del mundo*: provee acceso a la representación digital del mundo.
- *Presentación*: Sistema responsable de la presentación de los resultados. Incluye renderizado 2D/3D y salida de audio o táctil.
- *Contexto*: provee a la aplicación con contexto tanto del estado como de la situación del usuario.
- *Interacción*: representa la interacción del usuario con el sistema de RA.

De acuerdo a lo antes mencionado, en la figura 2 se muestra la vista de componentes del modelo referencial propuesto por [Mac05]:

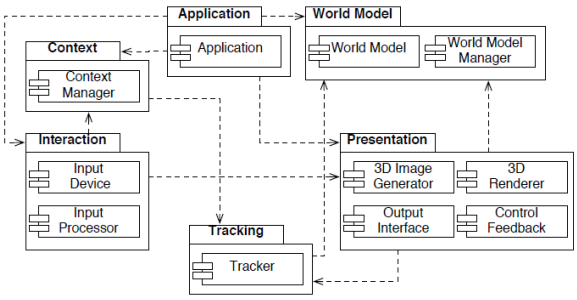


Figura 2

Cada subsistema está formando por varios componentes. El modelo de referencia describe a nivel abstracto los componentes generales y la estructura de los sistemas de realidad aumentada. Sin embargo, dependiendo de los requisitos funcionales de un sistema particular, algunos de los componentes pueden dejarse afuera [Mac04].

Pruebas de concepto y desarrollo de subsistemas

Con el objetivo de ir entendiendo y experimentando con la librería OpenCV, se realizaron pruebas de concepto, entendidas como subproyectos del proyecto principal ya que sus resultados aportaban componentes y conocimientos para el desarrollo del framework. A continuación se describen brevemente las más importantes:

Reconocimiento de Marcadores

Una de las técnicas más difundidas para la determinación de la pose (y en general para situar objetos 3D en el mundo real) es la que está basada en el reconocimiento de marcadores. Éstas son figuras que contienen rasgos geométricos fácilmente detectables (bordes, esquinas, etc.) que, en su forma canónica, ya están almacenadas en el sistema, por lo que a partir de detectar su posición en la escena real, se pueden calcular las variaciones respecto de la forma almacenada y las transformaciones geométricas necesarias para representar esta variación o lo que es lo mismo, la relación entre el sistema de coordenadas centrado en la cámara y el del mundo real. De esta manera se determina la pose de la cámara.

Los pasos involucrados en el reconocimiento de marcas son complejos y abarcan [Mit14b]:

- Captura y conversión del frame de entrada
- Transformación a escala de grises
- Binarización según un umbral
- Detección del contorno
- Determinación de bordes y vértices
- Cálculo de las coordenadas de la cámara relativas a la marca

Sistema de ayuda de navegación y obtención de información de aulas en entornos delimitados

La aplicación se concibió como un sistema de ayuda para la navegación y obtención de información académica dentro del edificio de la Facultad de Tecnología Informática. Se colocaron diferentes códigos QR para cada aula y laboratorio de la facultad y se desarrolló una

aplicación móvil para smartphones funcionando bajo sistema Android.

En la primera etapa se implementó una aplicación que permitía identificar los espacios de trabajo enfocando con el smartphone el frente del aula o laboratorio de forma que el lector de códigos QR pudiera recuperar la información codificada en el marcador y presentarla en la pantalla del dispositivo como información de texto. Dicha información que se presentaba estaba formada por: información de laboratorios/aulas, asignaturas que se dictan en los mismos, horarios y profesores a cargo (figura 3).

Figura 3

Las última etapa, aún en desarrollo, permitiría poner un objetivo espacial que se quiere alcanzar (aula o laboratorio), y la lectura de los marcadores que el navegante fuera encontrando en el entorno delimitado permitiría generar y superponer a la imagen captada por el dispositivo información de dirección a tomar (flechas) para llegar al destino.

Reconocimiento de objetos utilizando Deep Learning

El subproyecto, actualmente en curso, consiste en el desarrollo de un sistema de reconocimiento de objetos reciclables basado en Redes Neuronales Convolucionales [Ant17]. Este tipo de redes son muy efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones. Al momento de redactarse este artículo se está trabajando en la definición de algunas características de la arquitectura y se están reuniendo datos gráficos para armar una base de imágenes para el entrenamiento de la red. La función de la red será la de clasificar un objeto doméstico en alguna de 6 categorías objetos (p.ej.: envases). Este proyecto está sustentado por el trabajo de Tesina Final que está desarrollando una alumna del grupo de investigación del Laboratorio.

LINEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Diseño y desarrollo de un framework de RA y VA multipropósito (aplicado a diferentes disciplinas)
- Desarrollo de aplicaciones de RA y VA utilizando el framework propuesto
- Investigación de técnicas de Deep Learning aplicables al procesamiento de imágenes y clasificación de objetos utilizando como soporte el framework Keras [Ant17][Ker18]

RESULTADOS Y OBJETIVOS

A continuación se detallan los resultados obtenidos y los objetivos generales del proyecto:

- Se ha conformado un grupo de investigación y un Laboratorio dedicados al Procesamiento de Imágenes y Visión Artificial de la FITI en UB.
- El proyecto principal apunta a la creación de herramientas integradas en un entorno informático (framework) que facilite el desarrollo de aplicaciones de RA y VA.
- Dentro de este proyecto marco se estudian y exploran distintas tecnologías y

herramientas para entrenamiento de los sistemas en el reconocimiento de imágenes, generación de objetos. Estas pruebas de concepto fueron desarrolladas con el fin de ir materializando los conocimientos teóricos que se iban adquiriendo.

- Consistente con la potencialidad de aplicaciones basadas en el uso de Redes Neuronales Convolucionales y la disponibilidad de bases de imágenes rotuladas, aptas para el entrenamiento de estas redes mediante Deep Learning, se ha decidido dedicar una parte importante del proyecto a profundizar en el área de Reconocimiento de Objetos y desarrollar arquitecturas de redes multicapa capaces de clasificar una imagen entre un número limitado de categorías de objetos**.**
- Curso de Posgrado en Interfaces Avanzadas: Del mundo Real al Virtual. Departamento de Posgrado y Educación Continua. Facultad de Ingeniería y Tecnología Informática. Universidad de Belgrano.
- En el mes de noviembre del 2018 se presentó en el Instituto Inmaculada Concepción las capacidades de esta tecnología y los logros alcanzados en el laboratorio.
- Capacitación en el uso de procesamientos de imágenes, visión artificial y aprendizaje profundo a la empresa CyS en el marco del proyecto denominado Mapas Mentales financiado por el FONTAR.

FORMACIÓN DE RECURSOS HUMANOS

Uno de los principales objetivos de este proyecto de investigación y de la creación del Laboratorio de Procesamiento de imágenes y Visión Artificial es la formación de recursos en las temáticas de Realidad Aumentada y Visión Artificial enfocándose particularmente en Redes Neuronales Convolucionales para la clasificación de imágenes.

Durante el período de trabajo se terminó de elaborar una tesis de doctorado y se avanzó en la redacción de una tesina de grado:

- Alejandro Mitaritonna. “Realidad Aumentada para la Identificación de Objetivos Militares” (culminación de tesis de doctorado en la UNLP. Director: Dra. María José Abásolo; Codirector: Dr. Francisco Montero Simarro)
- Silvana Olmedo. “Sistema de Reconocimiento de objetos reciclables” (tesina de grado en Ingeniería Informática en la FITI-UB. Tutor: Mag. Alejandro Mitaritonna)

Asimismo se han capacitado a los alumnos que participan en el proyecto en el marco de esta línea de investigación: Carolina Páez, Florencia Vela, Tomás Poeta, Martín Lorenzo y Leonardo Inza.

Por último, nuestra intención durante el presente año es realizar demostraciones y presentaciones en Colegios u otros establecimientos de estudios e inclusive de investigación del uso de esta tecnología y los avances logrados en el Laboratorio.

REFERENCIAS

[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.

[Ant17] Antona Cortés, C. Herramientas modernas en redes neuronales: La librería Keras. Trabajo Final de Grado. Univ. Autónoma de Madrid, 2017

[Azu97] R. T. Azuma. A Survey of Augmented Reality. Teleoperators and Virtual Environments 6, 4, 355-385, 1997.

[Ker18] Keras: The Python Deep Learning library. https://keras.io/

[Mac04] A. MacWilliams, T. Reicher, G. Klinker y B. Bruegge, Design Patterns for Augmented Reality Systems, Institut für Informatik, Technische Universität München: In Proc. International Workshop Exploring the Design and Engineering of Mixed Reality Systems - MIXER 2004, 2004.

[Mac05] A. MacWilliams, A Decentralized Adaptive Architecture for Ubiquitous Augmented Re-ality System, Institut fur Informatik der Technischen Universitat Munchen, 2005.

[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.

[Mit14a] Mitaritonna, A., Pandolfo, L., Yokhdar D., Esteves, C. "RAIOM. Introducción a la arquitectura del framework de Realidad Aumentada". Escuela y Workshop Argentino en Ciencias de las Imágenes. ECIMAG 2014. Facultad de Ingeniería, UBA, 11 al 15 Agosto 2014.

[Mit14b] Mitaritonna, A., Pandolfo, L., Yokhdar D., Esteves, C. "RAIOM. Introducción a los algoritmos de visión por computador". Escuela y Workshop Argentino en Ciencias de las Imágenes. ECIMAG 2014. Facultad de Ingeniería, UBA, 11 al 15 Agosto 2014.

[Mit15] Mitaritonna, A.; Abásolo Guerrero, M. J. (2015) Improving Situational Awareness in Military Operations using Augmented Reality. Proceedings of WSCG 2015. ISBN Nº:978-80-86943-72-5, 2015

[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.

[Mit18] Mitaritonna, A.; Lestani, Juan;. Tarulla, Francisco; Poeta, Tomás; Olmedo, Silvana; Paez, Carolina; Lorenzo, Martín; Vela, Florencia; Inza, Leonardo. "Realidad Aumentada y Visión Artificial. Framework multipropósito". 20º WICC, UNNE, Argentina, 2018.

Presencia e Inmersión en Realidad Virtual: factores que maximizan la sensación de estar dentro del ambiente virtual

Matías N. Selzer, Martín L. Larrea y Silvia M. Castro

Departamento de Ciencias e Ingeniería de la Computación,
Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica, (UNS-CIC Prov. de Buenos Aires)
{matias.selzer, mll, smc }@cs.uns.edu.ar

Resumen

La Realidad Virtual es una tecnología que intenta sumergir a los usuarios en un mundo virtual generado por computadora. Aunque suele pensarse que el realismo gráfico es lo más importante, existen además otros factores que influyen en las sensaciones de presencia e inmersión percibidas, como el tipo de sonido, en campo de visión, la resolución del display, etc. La definición e influencia de dichos factores ha sido muy discutida en la literatura. Esta línea de investigación tiene como objetivo relevar dichos factores y analizar su impacto y relación con distintas técnicas computacionales y estadísticas. Se busca desarrollar una métrica que prediga el nivel de presencia e inmersión percibida por cualquier sistema de Realidad Virtual dado. Esta métrica y la importancia de cada factor brindará una guía para poder seleccionar el sistema y técnicas de Realidad Virtual que resulten más eficientes para la aplicación requerida.

Palabras clave: *Realidad Virtual, Interacción Humano Computadora en Realidad Virtual, Presencia e Inmersión*

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, becarios doctorales y alumnos de grado.

1. Introducción

La Realidad Virtual (RV) es una rama de las ciencias de la computación que promete sumergir al usuario en un mundo generado por computadora, al mismo tiempo que intenta desvanecer la línea entre lo real y lo ficticio. El avance tecnológico de los últimos años ha facilitado la posibilidad de generar y experimentar entornos virtuales y simulaciones cada vez más realistas. El mercado se encuentra actualmente inundado con juegos y simulaciones para entretener, motivar y divertir a los usuarios. Por el otro lado, el ámbito científico y comercial se enfoca principalmente en la creación de aplicaciones y sistemas muy específicos en áreas como la medicina ([1, 2, 3]), la arqueología ([4, 5]), el entrenamiento militar ([6]), o diferentes tipos de simulaciones ([7]).

En RV, la presencia se define como "la sensación de estar ahí", es decir, la sensación de que uno de verdad pertenece a el mundo virtual que actualmente está experimentando ([8]). La inmersión, por el otro lado, está más relacionada al hardware utilizado y se define como "qué tanto de nuestros sentidos está comprendido por el sistema de RV?" ([9]). Por este motivo, la

presencia es una medida subjetiva que depende de la percepción de cada usuario; mientras que la inmersión es una medida objetiva que depende del sistema de RV que se esté utilizando. Se suele pensar que en un sistema de RV es fundamental maximizar el nivel de realismo gráfico para obtener la mayor sensación de presencia. Sin embargo, está demostrado que esto no es necesario ([10]).

Desde los inicios de la RV, muchos investigadores han tratado de definir qué aspectos o factores de un sistema de RV verdaderamente influyen en el nivel de presencia percibida. Algunos de los más relevantes incluyen la resolución de la pantalla, la calidad gráfica, el nivel de sonido, entre otras ([11, 12, 13]). Sin embargo, existe mucha discusión entre la verdadera influencia de dichos factores e incluso la relación entre ellos. Lo que es más, la mayoría de dichos trabajos no son actuales, por lo que sus experimentos no aprovechan las posibilidades de las tecnologías modernas.

Por este motivo, es de vital importancia realizar un estudio exhaustivo de todos aquellos factores que pueden influir en la sensación de presencia, y por lo tanto también en la sensación de inmersión producida por cualquier sistema de RV. Dicho análisis revelará los factores más relevantes que deben tenerse en cuenta a la hora de crear un nuevo sistema o a la hora de seleccionar uno de los sistemas existentes. Como la presencia y la inmersión son medidas generales de la percepción de los usuarios, este estudio beneficia cualquier área de aplicación que desee incorporar o complementarse con la tecnología de RV.

2. Líneas de Investigación, Desarrollo e Innovación

Actualmente esta línea de investigación se enfoca en el análisis y relevamiento de todos aquellos factores que influyen en la sensación de presencia e inmersión en un sistema de RV. La cantidad de combinaciones de todos los componentes de dichos factores es muy extensa y no existe una forma directa de calcular el nivel de presencia o inmersión que produce cada una de ellas. Por este motivo, la participación de un usuario es pertinente para poder calificar cada una de dichas combinaciones. Aunque no será posible calificar todas las combinaciones, con un gran número basta para realizar un análisis. A partir de esto, nuestra línea de investigación trata de hallar una métrica que relacione aquellos factores que influyen mayormente en la determinación de estos niveles de presencia e inmersión. Con esta métrica se puede entonces predecir el nivel de presencia o inmersión que generaría cualquier sistema de RV dado.

A partir de las combinaciones de factores generadas con su correspondiente calificación del usuario, se forma un dataset que puede utilizarse para realizar diversos análisis.

- **Predictor de presencia e inmersión:** dado un determinado sistema de RV, resulta importante predecir el grado máximo de presencia e inmersión que éste produce. Esta línea trabaja en conjunto con técnicas de Machine Learning para la generación de dicho predictor a partir del dataset ya construido.
- **Importancia y optimización de los factores:** como la cantidad de factores

que influyen en el nivel de presencia e inmersión de los sistemas de RV es muy grande (al menos 30), resulta difícil optimizar manualmente cada uno de ellos cuando se desea desarrollar un nuevo sistema. Por este motivo, es importante identificar aquellos factores que más influyen en el nivel de presencia e inmersión final. Al reducir considerablemente la cantidad de factores, el diseñador de un nuevo sistema puede enfocarse en la optimización de los mismos, mejorando así el grado de presencia e inmersión del sistema.

- **Análisis estadísticos:** a partir del dataset construido se pueden comparar los factores entre sí para determinar si existe alguna correlación entre los mismos. De esta forma se puede calcular la interacción entre ellos y la influencia que tienen en el valor final de presencia e inmersión.

3. Resultados y Objetivos

Sobre los ejes presentados se han obtenido resultados parciales. Actualmente se está desarrollando una aplicación de RV que genere aleatoriamente un escenario distinto en función de cada uno de los factores identificados. Por ejemplo, un escenario tendrá determinada resolución de pantalla, nivel gráfico, tipo de audio, tipo de tracking, etc.; mientras que otro escenario generado tendrá factores completamente distintos. La aplicación permitirá a los usuarios experimentar un escenario generado durante un tiempo determinado para que finalmente dichos usuarios califiquen el escenario según el nivel de presencia experimentado. Los usuarios podrán realizar esto todas las veces que deseen y cada prueba constituirá una nueva muestra del dataset. Esperamos que en sistema esté online todo el tiempo posible para que el dataset siga creciendo, lo que ayudará a obtener mejores predictores y análisis estadísticos entre las variables.

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:

Tesis Concluidas

"Interacción Humano Computadora en Ambientes Virtuales", tesis de posgrado para el Magíster en Ciencias de la Computación. Alumno: Matías Selzer. Director: Martín Larrea.

Tesis en Desarrollo

"Métricas de Inmersión para sistemas de Realidad Virtual", tesis de posgrado para Doctor en Ciencias de la Computación. Alumno: Matías Selzer. Director: Silvia M. Castro. Co-Director: Martín L. Larrea.

"La Realidad Virtual como Herramienta para el Desarrollo Arquitectónico", tesis de grado de Ingeniería en Sistemas de Información. Alumnos: Facundo Reissing, Sebastián Vicente. Director: Martín L. Larrea. Colaborador: Matías N. Selzer.

5. Bibliografía

1. 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*.

2. 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*.
3. Gorini, A., & Riva, G. (2014). Virtual reality in anxiety disorders: the past and the future. *Expert Review of Neurotherapeutics*.
4. Lynch, J., & Corrado, G. (2014). Arqueología virtual aplicada al sitio Villavil, Catamarca, Argentina Virtual Archaeology applied to the site Villavil, Catamarca, Argentina.
5. 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.
6. Carroll, J. M. (Ed.). (2003). *HCI models, theories, and frameworks: Toward a multidisciplinary science*. Morgan Kaufmann.
7. 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*.
8. Bob G Witmer and Michael J Singer. Measuring presence in virtual environments: A presence questionnaire.Presence, 7(3):225–240, 1998.
9. Mel Slater. Measuring presence: A response to the witmer and singer presence questionnaire.Presence, 8(5):560–565, 1999.
10. Rheingold, H. (1991). Virtual reality: exploring the brave new technologies. Simon & Schuster Adult Publishing Group.
11. Khanna, P., Yu, I., Mortensen, J., and Slater, M. (2006). Presence in response to dynamic visual realism: A preliminary report of an experiment study. In Proceedings of the ACM Symposium on Virtual Reality Software and Technology, VRST ’06, pages 364–367,New York, NY, USA. ACM.
12. Jsselsteijn, W., Ridder, H. d., Freeman, J., Avons, S. E., and Bouwhuis,D. (2001). Effects of stereoscopic presentation, image motion, and screen size on subjective and objective corroborative measures of presence.Presence: Teleoperators and Virtual Environments, 10(3):298–311.
13. Skarbez, R., Brooks, Jr., F. P., and Whitton, M. C. (2017a). A survey of presence and related concepts.ACM Comput. Surv., 50(6):96:1–96:39.

Segmentación de imágenes de partículas de arena utilizadas para fracking

Mg. Ing. Carlos Gustavo Rodriguez Medina [1], Dr. Ing. Oscar Daniel Chuk [2],
Ing. Regina Bertero, Lic. Adriana Luna, Ing. Enrique Núñez. Ing. Pablo Trigo

Instituto de Investigaciones Mineras. Facultad de Ingeniería.
Universidad Nacional de San Juan

Av. Libertador Gral. San Martin 1109 oeste. San Juan
0264-4211700 (int. 285 [1], int. 389 [1,2])

grodriguez@unsj.edu.ar [1], dchuk@unsj.edu.ar [2]

RESUMEN

El presente trabajo se enmarca en el área de procesamiento digital de imágenes. La línea de investigación se orienta en poder determinar la calidad de arenas utilizadas en procesos de fracturación hidráulica (también conocida como *fracking*) para la extracción de petróleo de manera no convencional. Dicha determinación de calidad se realiza mediante la medición de características, tales como la esfericidad y la redondez de las partículas de arena, a través de técnicas de procesamiento de imágenes en una PC.

Se ha tenido que invertir un importante esfuerzo en poder segmentar, de manera apropiada, las imágenes de interés (partículas o granos de arena) para luego poder determinar las características antes mencionadas, indicadoras de calidad.

Es por ello que se ha trabajado en la elaboración de un algoritmo que aplica la técnica de Divisoria de Aguas, mediante la Transformada de Watershed, obteniéndose así una adecuada segmentación para los fines perseguidos.

Si bien la utilización de la técnica de Divisoria de Aguas se ha encontrado solo en publicaciones actuales referidas al procesamiento de imágenes biomédicas, para este caso se hace uso en una aplicación con fines industriales.

Palabras clave: Segmentación de imágenes, Divisoria de aguas, Watershed, Partículas de arena, Fracking.

CONTEXTO

El desarrollo del trabajo que se expone a continuación se lleva a cabo en el marco de un Proyecto de Investigación y Creación (PIC), titulado *"Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad"*. El mismo es parte de una convocatoria más amplia de Proyectos, tal como los PDTS (Proyectos de Desarrollo Tecnológico y Social), FODO (Formación Docente), y PROJOVI (Proyectos de Jóvenes Investigadores).

El mencionado Proyecto tiene un plazo de ejecución de dos años, y se inició en 2018.

La financiación es llevada a cabo con fondos correspondiente al presupuesto de la Universidad Nacional de San Juan (UNSJ), como así también por parte del Gobierno Provincial. La evaluación de propuestas iniciales de proyectos, como así también la evaluación final, es llevada a cabo mediante evaluadores externos a la UNSJ.

1. INTRODUCCIÓN

El proceso de Fracking o el Fracturamiento Hidráulico, es el proceso utilizado en la extracción de hidrocarburos de manera no convencional, como es el caso del yacimiento de Vaca Muerta, en la provincia de Neuquén. En dicho proceso se aplica presión hidráulica a la roca reservorio de hidrocarburo, hasta que se produce una falla o fractura de la misma mediante la inyección a alta presión de fluido viscoso, y se mantiene abierta con un agente apuntalante (tal como arena natural o sintética) una vez que se haya liberado la presión de inyección.

El agente de sostén, agente apuntalante o propante es un material usado para evitar que la fractura generada en la roca durante el Fracturamiento Hidráulico se cierre completamente, una vez terminado el bombeo y garantizar una buena conductividad al canal recién creado.

Con la finalidad de garantizar la capacidad de las arenas con destino al Fracking para actuar como apuntalantes o propantes, se reconocen como base las normas de referencia del *American Petroleum Institute* (API). Esta norma recomienda una serie de ensayos a realizar sobre las arenas con el fin de conocer la calidad del material a utilizarse en el Fracking. Entre Los ensayos a realizarse se encuentran los de Redondez y Esfericidad de los granos de arena.

Los ensayos de Esfericidad y Redondez padecen de una fuerte dependencia del criterio del observador, dado que el procedimiento se realiza mediante la utilización de un microscopio y las mediciones de esfericidad y redondez se realizan por comparación respecto a tablas que contienen una serie de formas provistas por la norma. Esto deriva en resultados sustancialmente disímiles que invalidan el procedimiento, aunque es el que se sigue usando por norma.

El uso de técnicas de Visión Artificial se presenta como la alternativa tecnológica que puede permitir obtener medidas objetivas no dependientes del criterio de un operador. [1].

Al operar con imágenes correspondientes a muestras de arenas, los elementos de interés se esparcen sobre una superficie contrastante (utilizada como fondo), teniendo la precaución de evitar el amontonamiento de las partículas, ya que cuando se captura la imagen y al quedar partículas unidas unas a otras un proceso de segmentación común (como una umbralización) no es suficiente. Por mayor cuidado que se tenga, generalmente se produce algún efecto de amontonamiento o aglomeración de las partículas de arena, es prácticamente inevitable.

Por tal motivo, resulta de interés encontrar algún mecanismo para salvar dicha situación, dado que para estos casos la segmentación no arroja resultados satisfactorios, ya que se suelen obtener elementos que no representan las partículas de arena reales.

En el marco de este trabajo, se han probado distintos algoritmos de segmentación, pero con el que mejores resultados se ha podido obtener (en cuanto a lograr la división de partículas unidas a nivel imagen) ha sido mediante el algoritmo de Divisoria de aguas.

Esto, de alguna manera confirma lo expresado por [2], que indica que la partición de una imagen en regiones permite obtener información que busca el observador mediante una herramienta morfológica poderosa denominada transformada de Watershed, también conocida como Divisoria de Aguas. Dichos autores manifiestan que tal transformada es fácilmente adaptable a los diferentes tipos de imágenes y es capaz de distinguir objetos sumamente complejos que no pueden ser analizados correctamente mediante algoritmos convencionales.

Watershed es una técnica de segmentación basada en morfología matemática, que permite extraer las fronteras de las regiones que hay en una imagen. A la vez, se considera una técnica de segmentación basada en regiones, debido a que clasifica los pixeles según su proximidad espacial, el gradiente de sus niveles de gris, y la homogeneidad de sus texturas. Por ello se toma como una técnica de detección de contornos y crecimiento de regiones al mismo tiempo [2] [3].

A través de Watershed se puede considerar una imagen en escala de gris como la imagen topográfica de un relieve terrestre; en donde a cada pixel se le asocia como valor de altura su nivel de gris correspondiente. En este sentido,

se puede pensar que las intensidades de gris de mayor amplitud se corresponden con montañas, mientras que las intensidades de menor valor se corresponden con valles y ríos [3].

La técnica además incorpora un proceso de inundación de los valles, desde los niveles más bajos de altura (valores mínimos locales, que constituyen cuencas de inundación rodeadas por cadenas montañosas), hasta los niveles más altos.

Las zonas de baja intensidad de gris también se conocen como vasijas (basins en Inglés), por donde fluirá el agua e inundará toda la topografía de la imagen; es decir, el agua fluirá en cada una de las vasijas identificadas. El proceso de inundación continúa hasta que las aguas de cuencas o vasijas contiguas se unen, formando líneas de unión que representarán las fronteras de regiones homogéneas, y constituyen el resultado de la segmentación. Lo descripto se presenta en la figura 1.

Las líneas divisorias de aguas forman los contornos de los objetos de la imagen que se quiere analizar, que darán lugar a la segmentación de la imagen, y que corresponden además a zonas de elevada intensidad de gris.

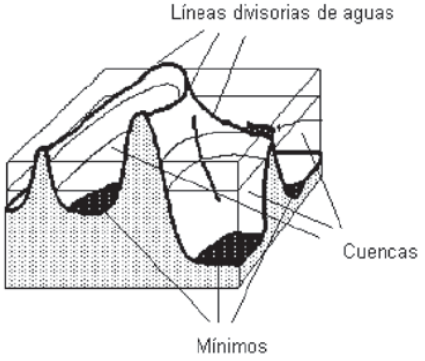


Figura 1. Watershed: Líneas divisorias, mínimos y vasijas o cuencas. Fuente [La Serna y García 2011].

Se debe tener especial cuidado durante la aplicación del algoritmo de Watershed, dado que en imágenes no homogéneas y/o embebidas en ruido se genera un gran número de mínimos locales, produciéndose sobresegmentación en pequeñas regiones, en donde muchas de las regiones generadas no son importantes dentro de la imagen, o no representan objetos existentes en la imagen original. Esto genera errores en las mediciones o propiedades de los elementos de interés de la imagen [2].

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

En el Laboratorio de Control Automático del Instituto de Investigaciones Mineras (Facultad de Ingeniería de la Universidad Nacional de San Juan) se realiza investigación y desarrollo respecto a la automatización de procesos de molienda de minerales hace casi dos décadas. Hace algunos años se viene desarrollando el Área de Procesamiento de Imágenes.

Actualmente se está trabajando en la caracterización de arenas especiales (forma, color, tamaños, clasificación del tipo de material, etc.) mediante el procesamiento de imágenes para ser utilizadas en la extracción de petróleo mediante el método de Fraking.

Otra de las líneas de investigación y desarrollo del grupo de trabajo, está orientado al de optimización. Concretamente, se está trabajando en la optimización multiobjetivo, del tipo predictivo, para la explotación de minas.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Para el procesamiento digital de las imágenes, se utiliza MATLAB. Este es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio (lenguaje M). Está disponible para las plataformas Unix, Windows y Apple Mac OS X. [4].

La ventaja de operar imágenes con Matlab, es que las mismas se cargan al entorno de trabajo en forma de matrices, y dado que este software está orientado a aplicar toda su potencialidad de cálculo a las matrices, es que resulta una herramienta potente para la manipulación y procesamiento matemático; además de poder trabajarse digitalmente las

imágenes a través del Toolbox de Procesamiento de imágenes con el que cuenta (por ej. aplicar filtros, mascaras, transformadas, etc.).
En la figura 2 se expone el algoritmo elaborado, el que permite segmentar las partículas de arena de manera adecuada, eliminado el problema antes descripto referido al amontonamiento (aglutinación) de granos de arena, como si estuvieran pegados unos a otros.

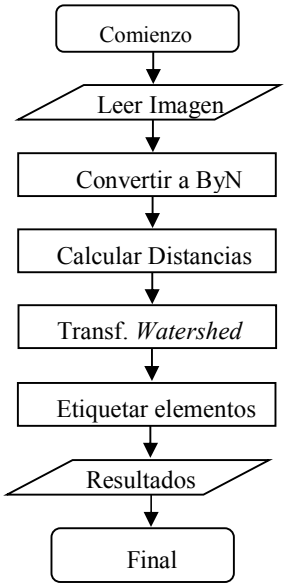
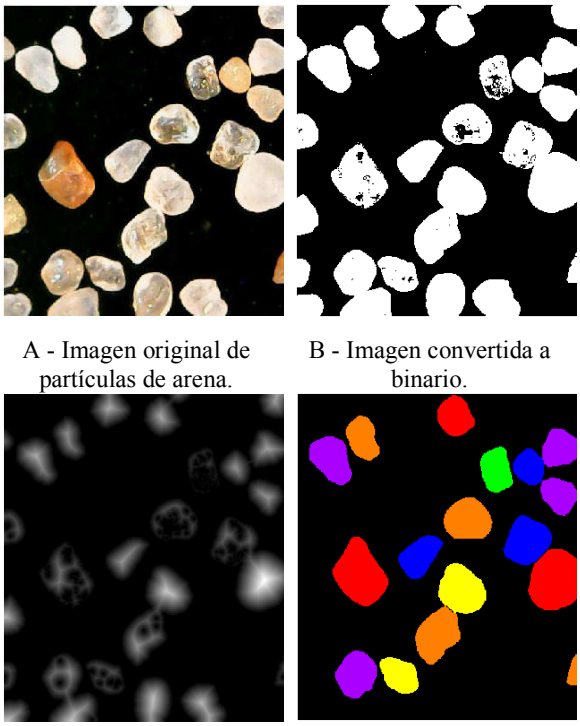


Figura 2. Algoritmo para segmentar imágenes de partículas de arena, eliminando problema de aglutinación.

Como primer paso del algoritmo, se carga en el espacio de trabajo de Matlab la imagen de interés a ser procesada. Se dispone de una función específica para ello, "imread()", la que permite que la imagen digital, almacenada en forma de un archivo en el disco de la PC, se cargue en memoria como una matriz, de NxMxR. Donde N hace referencia a la cantidad de filas, mientras que M a la cantidad de columnas y R al número de planos de color de la imagen, es decir a los planos de color R(rojo), G(verde) o B(azul) de una imagen de 24 bpp color [5].
Luego la imagen se convierte en binaria, o blanco y negro, para así poder aplicarle una transformada de distancia. Esto es a los fines de poder encontrar elementos que definirán las regiones o vasijas para la posterior aplicación de Watershed [2] [3].
La transformada de distancia de una imagen binaria es la distancia de cada pixel al pixel más cercano distinto de cero [6]. Suele utilizarse en operaciones morfológicas tales como la erosión, dilatación y en la segmentación.
Se utiliza en forma conjunta con la transformada de Watershed para realizar la segmentación, para dividir formas convexas que se tocan entre si y poder considerarlas como elementos separados.
Para el cálculo de distancia, se pueden utilizar distintas métricas, haciendo uso de métodos tales como: Euclídea, City block, Chessboard y Cuasi-euclídea.
Para el caso del presente trabajo, se utiliza Watershed en conjunto con la transformada de distancia de los pixeles de la imagen. Se debe tener en cuenta que la utilización de esta técnica puede llegar a producir algún nivel de sobresegmentación en la imagen de resultado. Es decir, pueden aparecer una mayor cantidad de líneas divisorias que la cantidad de elementos a segmentar, subdividiéndose dichos elementos de interés, resultado esto en un efecto no deseado.
En la figura 3, se exponen una serie de imágenes, las que corresponden a un ejemplo concreto de aplicación del algoritmo elaborado. La primera de ellas, figura 3-A, se refiere a la imagen original (sin ningún procesamiento) de una muestra de partículas de arena, utilizada como entrada al algoritmo desarrollado. La figura 3-B, corresponde a la imagen convertida a binario. Para el caso de la figura 3-C, se presenta el resultado de la transformada de distancia de la imagen binaria, para luego aplicarle la transformada de Watershed y así obtener la imagen expuesta en la figura 3-D.
La imagen correspondiente a la figura 3-D, es la de mayor interés como resultado obtenido. En ella se aprecian las partículas de arena segmentadas correctamente, aun para los casos en los que se hallan partículas que se tocan entre sí. Nótese que se hace referencia a una segmentación correcta o adecuada, desde el punto de vista que no se presentan casos en

que el algoritmo confunda dos partículas pegadas como si fuera una sola, o también algún otro defecto, tal como el de sobresegmentación.



A - Imagen original de partículas de arena.

B - Imagen convertida a binario.

C- Imagen resultado de la transformada de distancia.

D - Imagen resultado de aplicar *Watershed*.

Figura 3. Resultados por etapas al aplicar el algoritmo de la figura 2.

Se aprecia la utilización de diferentes colores en las partículas de arena, esto es presentado así por el propio algoritmo elaborado, y es a los fines demostrativos de que el algoritmo las ha considerado como elementos diferentes, logrando así una segmentación adecuada. El algoritmo etiqueta cada uno de los elementos obtenidos (segmentados) para luego poder operar sobre ellos en forma independiente, y poder aplicar las métricas que correspondan.

Se han llevado a cabo diversas pruebas del algoritmo elaborado, y para las imágenes de interés utilizadas se han podido obtener resultados muy satisfactorios. También se sabe que la sobresegmentación es un problema latente en este tipo de técnicas, y que se deben buscar otras alternativas para tratar de eliminarla o minimizarla, en el caso de que aparezcan.

A partir de estos resultados obtenidos, se pueden aplicar distintas técnicas y algoritmos para poder llevar a cabo las mediciones de redondez y esfericidad de las partículas de arena, como así también cualquier otra métrica que permita estimar la calidad de una determinada arena para Fracking.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de investigación está conformado por Ingenieros Electrónicos, Electromecánicos, en Minas y Metalurgistas, con formación diversa en posgrado, tal como, Doctorado en sistemas de control, Maestría en informática, Especialización en gestión y vinculación tecnológica, etc.

Actualmente, una de las integrantes del equipo de trabajo, se encuentra desarrollando su Tesis de Doctorado en Ingeniería en procesamiento de minerales, siendo su director de Tesis también otro de los integrantes del grupo de investigación.

Además, como inicio en la carrera de investigador, tres de los actuales integrantes, comenzaron hace tres años participando de un proyecto PROJOVI (Proyecto de Jóvenes Investigadores), cuya finalidad es la de permitir que docentes, becarios, y alumnos de la Universidad, menores de 40 años, con o sin experiencia en investigación lleven a cabo dicha actividad. Esto ha permitido su inserción, desde el año 2018, en un equipo de trabajo de mayor tamaño y experiencia.

5. BIBLIOGRAFIA

[1] Rodriguez Medina, C. G.; Chuk, O. D.; Bertero R.; Luna, A.; Núñez, E.; Trigo, P. 2019. Visión artificial aplicada a la determinación de propiedades geométricas de arenas de fracturación para el análisis de calidad. Universidad Nacional de San Juan. Informe de Avance de Proyecto.

[2] Gonzales, M. A. y Ballerin, V. L. 2008. Segmentación de imágenes utilizando la transformada de Watershed: obtención de marcadores mediante lógica difusa. IEEE

Latin American Transaction, Vol. 6, N° 2. p 223 - 228.
[3] La Serna P., N. y Garcia H., N. 2011. Implementación del algoritmo de Watershed para el análisis de imágenes médicas. Revista de Investigación de Sistemas e Informática. Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos. RISI 8(2), p 67 - 74. ISSN 1816-3823.
[4] Mathworks. (1994-2018). "Image Processing Toolbox". https://es.mathworks.com/products/ image.html ?s_tid=srchtitle.
[5] Rodriguez Medina, C. G. y Navas, G. S. 2014. Aplicación del filtro de Canny en la esteganografía digital. XVI Workshop de Investigadores en Ciencias de la Computación - WICC 2014 Ushuaia. RedUNCI. p 806 - 811. ISBN 978-950-34-1084-4.
[6] Fiter, E. L. 2012. Descripción, comparación y ejemplos de uso de las funciones de la toolbox de procesado digital de imágenes de MATLAB. Proyecto fin de carrera. Escuela Universitaria de Ingeniería Técnica de Telecomunicaciones, Universidad Politécnica de Madrid.

Sistemas motivadores del Sujeto basados en Realidad Virtual y Realidad Aumentada

G. Rodríguez, N. Jofré, Y. Alvarado, M. Zúñiga,
M. Rosas, J. Fernandez, 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
{gbrodriguez, npasinetti, ymalvarado, mezuniga, mvrosas, 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 Alterna- tivas, 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 aplicación de la RV y la RA en sistemas de acceso al conocimiento, con el fin de lograr en el sujeto cambios motivacionales e incentivar procesos que implican un esfuerzo físico y/o cognitivo, permitiendo así 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 Nacio- nal de San Luis.

1. INTRODUCCIÓN

Actualmente, la civilización transita nuevos escenarios cotidianos en los que se pueden identificar tres aspectos importantes: la glo- balización, el permanente cambio del contexto y la valoración del conocimiento.

Los tiempos presentes, sujetos a cambios drásticos y exponenciales, han generado grandes transformaciones a nivel social, cul- tural, educativo, de salud, entre otros; razón por la cual es necesario que los educadores/capa- citadores actuales reconozcan, analicen y reflexionen sobre el entorno donde las tecnologías de la información y la comu- nicación actúan como mediadoras y moti- vadoras en la formación de las personas [1,2,3].

La elaboración de un entorno de motivación y satisfacción se vuelve un elemento prioritario

del proceso de enseñanza y aprendizaje. Toda institución con un alto nivel de motivación en sus estudiantes, cuenta también con altos niveles de motivación en sus docentes. Más allá de las teorías de motivación que se han ido desarrollando en el transcurso del tiempo es necesaria la incorporación de mecanismos que den lugar a un **proceso de cambio motivacional** [4].

De acuerdo a estudios recientes, las tecnologías emergentes como la RV y RA han sido utilizadas con el fin de motivar a los estudiantes mediante metodologías educativas innovadoras. Los resultados reportados revelan que es posible mejorar la experiencia educativa de los estudiantes para motivarlos y compro- meterlos aprovechando las características de estas tecnologías [5,6].

Según resultados publicados por [7,8,9], la RV y RA en la educación son un tema emergente y la investigación sobre estas tecnologías en educación está en su fase inicial. Entre los datos reportados hasta el momento cabe destacar el uso de la RV y RA principalmente en la educación superior (grado, 23,64%) y en la educación obligatoria (prima- ria, 16,36%; secundaria, 18,18%) [10,11].

No obstante, las aplicaciones educativas son sólo un caso particular, pues 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 esce- narios virtuales ficticios o insertar objetos virtuales en el espacio real, las ha convertido en herramientas muy útiles para presentar deter- minados contenidos bajo diferentes premisas tales como la educación y la salud [12,13].

El impacto económico y social de las tecnologías mencionadas es tal que la inversión para facilitar el acceso a contenidos virtuales como los de la RV y RA a gran escala en el primer trimestre de 2016 fue de 1.200 millones de dólares, y se espera que obtenga ganancias de 120.000 millones de dólares en 2020 [18].

El empleo de sistemas de RV y RA constituyen un nuevo enfoque para entrenar, tratar y educar a las personas reforzando el aprendi- zaje a través de plataformas multimodales [14,15,16,17].

En consecuencia, las realidades alternativas creadas por medio de la RV y RA pueden incentivar la generación de cambios motivacionales positivos logrando comprometer al individuo en el desarrollo de una actividad específica.

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 Y DESARROLLO

Desde hace un tiempo, en el ámbito educativo se están incentivando elementos como el juego digital para involucrar a los estudiantes, motivarlos a la acción y promover el aprendizaje y la resolución de problemas. Puntualmente, la **motivación** juega un papel fundamental dado que se utiliza para llamar la atención de los estudiantes con el fin de que dediquen tiempo a las actividades [19,20]. Sin embargo, captar el interés no es algo que se pueda lograr con patrones planos, rígidos y unificados, dado que cada una de las personas se motiva por razones diferentes. Se debe contemplar el hecho de que el interés se podrá conseguir, conservar o acrecentar, en función de factores tanto internos como externos. Estos factores se conocen como *motivación intrínseca* (MI): se evidencia cuando el individuo realiza una actividad por el simple placer de realizarla; y *motivación extrínseca* (ME): aparece cuando lo que atrae al individuo no es la acción que se realiza en sí, sino la posibilidad de recibir recompensas externas [21].

La Realidad Virtual y la Realidad Aumentada están en el foco de atención por parte de los

diferentes contextos formativos, dado que ofrecen una alternativa de enseñanza y aprendizaje a través de la comunicación humano-computadora. Dicho esto, se propone involucrar a los sistemas de RV y RA en dos líneas de investigación, de acuerdo a los enfoques de motivación mencionados.

- **Sistemas Intrínsecos:** consisten en sistemas donde el proceso de aprendizaje implica el "deseo" o las ganas de realizar una actividad y otorga como resultado un "placer". Este tipo de sistemas involucran a los sistemas basa- dos en juegos digitales conocidos como *sistemas de conflicto, metas y reglas*, cuyo objetivo es atraer al alumno a través de una trama y un trasfondo narrativo, estableciendo metas y obje- tivos y respetando ciertas reglas; y los sistemas de *inmersión y participación*, que proporcionan diversión al alumno mediante historias atractivas, mundos virtuales tridimensionales envolventes, la adaptación de la dificultad de los desafíos, etc. [19,22]. Algunos ejemplos de estos sistemas basados en RV son la plataforma *Savia* (creada por la multinacional *Indra*) y *Discovery VR*, entre otros. Por otra parte, entre los ejemplos de la RA se encuentran las aplicaciones de *Franky y Rocco*, *La sabana*, entre otras [23,24,25,26].

- **Sistemas Extrínsecos:** consisten en la acción de otorgar recompensas bajo el supuesto de que el aprendizaje tiene mejores resultados si se canaliza a través de premios. Estos sistemas, son acordes a los sistemas basados en juegos digitales conocidos como sistemas de *recompensa*, los cuales utilizan mecanismos de premiación al usuario; y los sistemas de *competencia*, que utilizan mecanismos de competencia "sana", no sólo entre los compañeros sino también frente a uno mismo [19,22]. Algunos ejemplos de estos sistemas son las aplicaciones de RV, *STRIVR* de Wallmart [28] y las de entrenamiento y educación de la empresa *PROVR* [29]; y la aplicación de RA, *ABC AR - Space Discovery* [27].

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 incentivar 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 fun-cionales 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 inter- venciones.

4. FORMACIÓN DE RECURSOS

Los trabajos realizados han permitido la definición de trabajos finales de carrera de la Licenciatura en Cs. de la Computación (4 finalizados), tesis de Especialización en Educación Superior (1 en ejecución), tesis de Maestría en Cs. de la Computación (2 en ejecución y 1 finalizada) y tesis de Doctorado en Ciencias de la Computación (1 en ejecu-ción).

Adicionalmente se ha obtenido una beca de iniciación a la investigación y una beca de perfeccionamiento en investigaciones 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. Gonida, Eleftheria N.; Lemos, Marina S. Motivation in Education at a Time of Global Change: Theory, Research, and Implications for Practice. En Motivation in Education at a Time of Global Change: Theory, Research, and Implications for Practice. Emerald Publishing Limited, p. 1-12, 2019.
5. Chen, Peng, et al. "A review of using Augmented Reality in Education from 2011 to 2016." Innovations in Smart Learning. Springer, Singapore, pp 13-18, 2017.
6. Martín-Gutiérrez, Jorge, et al. Virtual technologies trends in education. EURASIA Journal of Mathematics Science and Technology Education, vol. 13, no 2, p. 469-486, 2017.
7. H.-K. Wu, S. W.-Y. Lee, H.-Y. Chang, and J.-C. Liang, “Current status, opportunities and challenges of augmented reality in education,” Comput. Educ., vol. 62, pp. 41–49, 2013.
8. K.-H. Cheng and C.-C. Tsai, “Affordances of Augmented Reality in Science Learning: Suggestions for Future Research,” J. Sci. Educ. Technol., vol. 22, no. 4, pp. 449–462, Aug. 2012.
9. Freina, Laura; Ott, Michela. A Literature Review on Immersive Virtual Reality in Education: State Of The Art and Perspectives. eLearning & Software for Education, no 1, 2015.
10. Chen, Peng, et al. "A review of using Augmented Reality in Education from 2011 to 2016." Innovations in Smart Learning. Springer, Singapore, pp 13-18, 2017.
11. Hodgson, Paula, et al. Immersive Virtual Reality (IVR) in Higher Education: Development and Implementation. En Augmented Reality and Virtual Reality. Springer, Cham, p. 161-173, 2019.
12. T. Jung y M.C. Dieck. Augmented Reality and Virtual Reality: Empowering Human, Place and Business. Progressin. Springer International Publishing, 2017.
13. 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.
14. 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.
15. F. Pallavicini, N. Toniazzi, L. Argen-ton, 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.

16. 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.
17. 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.
18. Digi-Capital (2016). VR/AR investment in 2015 breaks out near $700 million. Available on: http://www.digi-capital.com/ (http://goo.gl/aD5pib).
19. Edutrends. Gamificación. Observatorio de Innovación Educativa del tecnológico de Monterrey, 2016.
20. Ruth S. Contreras Espinosa y José Luis Eguia. *Experiencias de gamificación en aulas*. InCom-UAB Publicacions. Institut de la Comunicació, Universitat Autónoma de Barcelona. ISBN 978-84-944171-6-0.
21. Kerr J. *Motivation and emotion in sport*. Hove, Psychology Press, 1997.
22. J. Torrente et al., "Fomentando la Creatividad: Creación de Escenarios de Aprendizaje Basados en Juegos," Una Guia para Profesores, pp. 1–46, 2011.
23. Savia. Plataforma educativa para personas con autismo; 2012. URL: http://www-.tecnologiasaccesibles.com/savia/.
24. https://www.discoveryvr.com/
25. https://anteludium.com/
26. https://apptk.es/apps/conoce-los-animales-la-sabana/.
27. https://appadvice.com/app/abc-ar-space-discovery/1378163550.
28. https://www.strivr.com/.
29. https://provr.io/.

Supervisión de sistemas mediante el uso de tecnologías de realidad aumentada en el contexto de industria 4.0

Martin Becerra[1], Jorge Ierache[1], María José Abasolo[2,3]
[1]Universidad Nacional de La Matanza, DIIT, Grupo de Realidad Aumentada Aplicada
Florencio Varela 1903, La Matanza, Buenos Aires, Argentina
jierache@unlam.edu.ar
[2] Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)

[3] Comisión de Investigaciones Científicas de la Pcia. de Bs. As.

Resumen

Las líneas de investigación aplicada tienen por objetivo generar capacidades de supervisión y explotación de sistemas industriales con el empleo de tecnologías de Realidad Aumentada. Esta línea se integra con la línea de desarrollo de sistemas basados en catálogos aumentados los cuales permiten la visualización de información añadida a partir de la captura de marcadores.

Palabras clave: Realidad Aumentada, Catálogo Aumentado, Sistemas de Supervisión, Industria 4.0, Sistemas basados en conocimiento.

Contexto

La línea de 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 Framework para la Generación de Templates en Sistemas de Catálogos de Realidad Aumentada.

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]. En los últimos años, la RA se ha expandido a diferentes campos de aplicación tales como educación, salud, turismo, marketing y entretenimiento.

El grupo de investigación Realidad Aumentada Aplicada tiene por objeto 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. 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].

Con el avance de la Industria 4.0 [8,] en particular el campo de RA, se han generado nuevos campos de empleo en el proceso de manufactura integrada por computadora donde se destacan las áreas de diseño y manufactura en la que podemos encontrar desarrollos que implementan sistemas de visualización de datos de sensores con tasas de refrescos cercanas a tiempo real [9, 10, 11, 12]. Hasta la fecha la RA ha sido reconocida en la industria como un apoyo interesante para aplicaciones de mantenimiento, montaje y reparación de maquinaria [13,14].

En la próxima sección se describirán las líneas de investigación presentadas [15] en las cuales se continúa trabajando: sistema de catálogos aumentados para crear marcadores de RA mediante el uso de templates e industria 4.0

orientada a la supervisión y explotación de sistemas industriales y de almacenamiento utilizando tecnologías de RA.

Líneas de Investigación, Desarrollo e Innovación

Sistema de Catálogos Aumentados

El objetivo de esta línea de investigación es el desarrollo del Sistema de Catálogos Virtuales Aumentados [14] [16], permite la generación, distribución y explotación de contenidos de RA por parte de usuarios finales. Un catálogo se compone 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. Mediante el uso de templates, se pretende simplificar 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.

Realidad Aumentada el Contexto de la Industria 4.0

El objetivo de esta línea de investigación aplicada se centra en el desarrollo de herramientas que apoyen a humanos en la toma de decisiones en la supervisión en el ámbito industrial, con el empleo de RA a fin de permitir aumentar las capacidades de los sistemas de monitoreo y control industriales. Esta línea integra tres áreas de conocimientos (figura 1):

-Sistemas industriales en el marco de Manufactura Integrada por Computadora (CIM)-Industria 4.0.

- Sistemas basados en conocimiento. (SBC)

- Realidad Aumentada (RA)

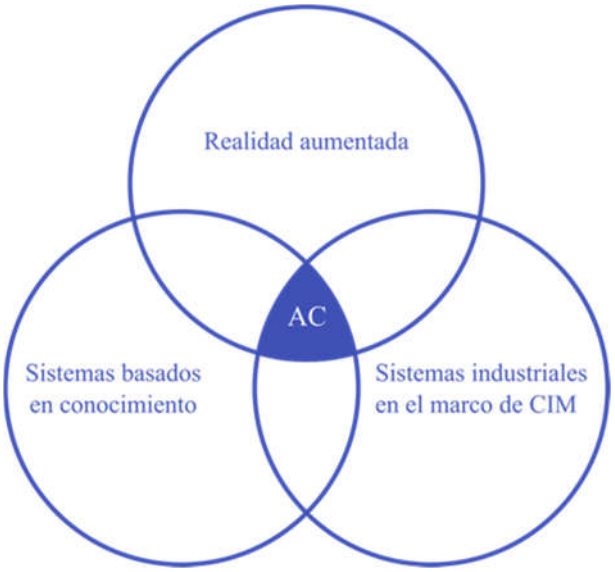


Figura 1 integración de Áreas de Conocimientos (AC) de la línea RA en el contexto de Industria 4.0

El objetivo es obtener como resultado un desarrollo de un prototipo que se integre a un sistema SCADA (Supervisory Control And Data Acquisition, en español Supervisión, Control y Adquisición de Datos) [17]. Se busca que el prototipo disponga de un sistema multiagente [18] para percibir eventos disparados por máquinas de una planta industrial, con el propósito de diagnosticar situaciones o estados de las máquinas y asistir 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 primera instancia se implementará para tablets, teléfonos móviles y gafas trasparentes. En segunda instancia se explorará el uso de comandos de voz para complementar y agilizar la interacción con los comandos de ejecución sobre el menú del sistema.

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 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 la supervisión y explotación, se desarrolló un prototipo experimental para la integración del sistema de gafas de cristal transparente Epson Moverio-300. Se tiene previsto explorar el uso de comandos de voz para complementar y agilizar la interacción con el prototipo, como así también explorar nuevas tecnologías de realidad virtual orientadas al adiestramiento en infraestructuras industriales

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por un investigador formado, dos investigadores en formación y un alumno del departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM, trabajando en el área de RA. Los investigadores en formación se encuentran realizando el Doctorado en Ciencias Informáticas en la UNLP, quienes se mencionan a continuación:

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), María José Abásolo (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] 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: 12 marzo 2019].

[9] Ong, S.K.; Huang, J.M. Structure design and analysis with integrated AR-FEA. CIRP Ann. Manuf. Tech. 2017, 66, 149–152.

[10] Huang, J.M.; Ong, S.K.; Nee, A.Y.C. Real-time finite element structural analysis in augmented reality. Adv. Eng. Softw. 2015, 87, 43–56.

[11] Huang, J.M.; Ong, S.K.; Nee, A.Y.C. Visualization and interaction of finite element analysis in augmented reality. Comput. Aided Des. 2017, 84, 1–14.

[12] Paulus, C.J.; Haouchine, N.; Cazier, D.; Cotin, S. Augmented reality during cutting and tearing of deformable objects. In Proceedings of the 2015 IEEE International Symposium on Mixed and Augmented Reality (ISMAR), Fukuoka, Japan, 29 September–3 October 2015.

[13] Azuma R. T., The Most Important Challenge Facing Augmented Reality, Presence Teleoperators Virtual Environment, vol. 25, n.o 3, pp. 234-238, dic. 2016.

[14] 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.

[15] Martin Becerra, Nahuel Mangiarua, Santiago Igarza, Jorge Ierache, María José Abasolo (2018) "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" XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018) ISBN 978-987-3619-27-4

[16] 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.

[17] Daneels A., Salter W., «What is SCADA?», 1999.

[18] Russell S., Norvig P., Inteligencia Artificial, un enfoque moderno, tercera edición 2010, cap. 2.

Técnicas biométricas: análisis de las técnicas actuales y nuevas tendencias

Lic. Fabiana A. Sánchez[1], Dra. Dana Urribarri[2], Dra. Silvia Castro[2]
[1] Departamento de Informática, Facultad de Ingeniería,
Universidad Nacional de la Patagonia San Juan Bosco
(UNPSJB, Prov. de Chubut)
fsanchez@ing.unp.edu.ar
[2] Laboratorio de I+D en Visualización y Computación Gráfica
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)
{dku, smc}@cs.uns.edu.ar

RESUMEN

En las tecnologías de la información, la biometría se refiere a la aplicación de técnicas matemáticas y estadísticas sobre los rasgos conductuales o físicos de un ser vivo para su autenticación, es decir, verificar su identidad. Existen muchas aplicaciones que requieren que la autenticación del usuario o la verificación de su identidad sea precisa, siendo la biometría la disciplina que se ocupa de la identificación automática o la verificación de la identidad de individuos basada en sus características fisiológicas y/o de comportamiento, resulta necesario comprender los indicadores biométricos completamente para poder implementar un sistema biométrico exitoso.

Palabras clave: Biometría, sistema biométrico, indicador biométrico

CONTEXTO

El presente trabajo se lleva a cabo entre docentes del Departamento de Informática de la Facultad de Ingeniería de la Universidad Nacional de la Patagonia San Juan Bosco (UNPSJB, Prov. de Chubut) y docentes investigadores del Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) del Departamento de Ciencias e Ingeniería de la Computación, de la Universidad Nacional del Sur; además, se enmarca dentro del proyecto en evaluación PGI 2019 *Detección de riesgo de Diabetes Mellitus tipo 2 a través de las huellas dactilares* presentado en la Universidad Nacional del Sur.

1. INTRODUCCIÓN

El origen de las biometrías se remonta 31.000 años atrás [10], existen cuevas que tienen sus paredes adornadas con pinturas creadas por los hombres prehistóricos que vivieron allí, donde se pueden encontrar numerosas impresiones de palmas de manos que fueron dejadas como las firmas de los autores. También existen registros donde figura que la biometría ya era utilizada en China desde al menos el siglo XIV. Joao de Barros fue un explorador y escritor que documentó que los comerciantes chinos registraban las huellas de la palma de las manos de los niños en papel con tinta. Sin embargo, la biometría no se puso en práctica en las culturas occidentales hasta finales del siglo XIX. En Occidente, la identificación confiaba simplemente en la memoria fotográfica [13] (denominada memoria de Eidetic) hasta que Alphonse

Bertillon, jefe del departamento fotográfico de la Policía de París, desarrolló en 1883 el sistema antropométrico, conocido más tarde como Bertillonage. Este fue el primer sistema preciso, ampliamente utilizado científicamente para identificar a criminales y convirtió a la biometría en un campo de estudio. Funcionaba midiendo de forma precisa ciertos largos y anchos de la cabeza y del cuerpo, así como registrando marcas individuales como tatuajes y cicatrices. El sistema de Bertillon fue adoptado extensamente en occidente hasta que aparecieron defectos en el sistema, principalmente problemas con métodos distintos de medidas y cambios de medida. Después de esto, las fuerzas policiales occidentales comenzaron a usar la huella dactilar, en esencia el mismo sistema visto en China cientos de años antes.

En estos últimos años la biometría ha crecido [12] desde usar simplemente la huella dactilar, a emplear muchos métodos distintos teniendo en cuenta varias medidas físicas y de comportamiento, como por ejemplo el andar, el tipeo o la voz. Tal es así, que el reconocimiento por biometrías es uno de los problemas mundialmente más investigados en la actualidad; los resultados de estas investigaciones han permitido alcanzar nuevos descubrimientos acerca del cuerpo humano y los patrones conductuales para expandir la lista de rasgos o características que son útiles para la identificación, e incluso el uso combinado de estos rasgos [8]. Por ejemplo, recientes descubrimientos han demostrado la posibilidad de usar la impresión de las orejas, ondas cerebrales, latidos del corazón y ADN como base para verificar la identidad [7]. Las aplicaciones de la biometría también han aumentado desde sólo identificación de los seres vivos hasta complejos sistemas de seguridad.

Existen muchas aplicaciones que requieren que la autenticación del usuario o la verificación de su identidad sea precisa. Dado que la biometría es la disciplina que se ocupa de la identificación automática o la verificación de la identidad de individuos basada en sus características fisiológicas y/o de comportamiento, resulta necesario comprender los indicadores biométricos completamente para poder implementar un sistema biométrico exitoso [11].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El objetivo de este trabajo es brindar una referencia extensiva de conceptos, definiciones, tecnologías actuales y nuevas tendencias en el área de los sistemas biométricos. Se cubrirán conceptos relacionados con las investigaciones, desarrollos y aplicaciones.

- Se efectuará una recopilación y un estudio de la bibliografía correspondiente a técnicas biométricas.
- Posteriormente se analizarán las técnicas biométricas actuales y nuevas tendencias para evaluar sus alcances y limitaciones.
- Se analizarán diferentes clasificaciones de técnicas biométricas y estrategias de comparación.
- Para concluir, se realizará un análisis comparativo de las técnicas biométricas con el fin de mejorar las clasificaciones existentes.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El resultado principal que se espera obtener es un análisis comparativo de las técnicas biométricas existentes. Para esto se realizará una recopilación exhaustiva de las técnicas

actuales, se analizará su aplicación, la relevancia y las nuevas tendencias a fin de detectar aquellas que se destaquen por su importancia.

4. FORMACIÓN DE RECURSOS HUMANOS

En lo que concierne a la formación de recursos humanos, este trabajo es el marco para la tesis de Maestría de la Lic. Fabiana Sánchez bajo la dirección de la Dra. Silvia Castro y la co-dirección de la Dra. Dana Urribarri. Adicionalmente, se incentiva la incorporación de alumnos que deseen realizar su tesina o trabajo final de carrera en temas relacionados.

5. BIBLIOGRAFÍA

[1] Biometric System Laboratory, University of Bologna, Italia, 1993

[2] BiDA Lab – ATVS Biometric Research Group, Universidad Autónoma de Madrid, España, 2016

[3] Mliki, H., & Hammami, M. (2018). Face analysis in video: face detection and tracking with pose estimation. International Journal of Biometrics, 10(2), 121.

[4] Harvey, J., Campbell, J., & Adler, A. (2018). Characterization of Biometric Template Aging in a Multiyear, Multivendor Longitudinal Fingerprint Matching Study. IEEE Transactions on Instrumentation and Measurement, PP, 1–9.

[5] Gawande, U., & Golhar, Y. (2018). Biometric security system: a rigorous review of unimodal and multimodal biometrics techniques, 10(2), 142–175.

[6] Eshwarappa, M., & Latte, M. (2018). Bimodal biometric person authentication system using speech and signature features. International Journal of Biometrics, 4(4), 147.

[7] Choudhury, B., Then, P., Issac, B., Raman, V., & Haldar, M. K. (2018). A Survey on Biometrics and Cancelable Biometrics Systems. International Journal of Image and Graphics, 18(01).

[8] Singh, A. (2018). Review on Multibiometrics: Classifications, Normalization and Fusion Levels. 2018 International Conference on Advances in Computing and Communication Engineering (ICACCE), 159–162.

[9] Blanco-Gonzalo, R., Lunerti, C., Sanchez-Reillo, R., & Guest, R. M. (2018). Biometrics: Accessibility challenge or opportunity? PLOS ONE, 13(3), 1–20.

[10] Sinjini, M., & Mikhail, G. (2017). Biometrics in a Data Driven World.

[11] Chahal, R. (2017). A comparative study of various biometric approaches. International Journal of Engineering Applied Sciences and Technology.

[12] Saeed, K. (2017). New directions in behavioral biometrics. CRC Press.

[13] Jain, A. K., Nandakumar, K., & Ross, A. (2016). 50 years of biometric research: Accomplishments, challenges, and opportunities. Pattern Recognition Letters, 79, 80–105.

[14] Nabil Ibrahim El-Sawalhi. (2015). Support Vector Machine Cost Estimation Model for Road Projects. Journal of Civil Engineering and Architecture, 9(9), 846–859.

[15] Li, S. Z., & Jain, A. K. (2015). Encyclopedia of Biometrics. Springer.

[16] Kisku, D. R., Gupta, P., & Sing, J. K. (2014). Advances in Biometrics for Secure Human Authentication and Recognition. In Biometrics. CRC Press.

[17] Yingzi, D. E. (2013). Biometrics: From Fiction to Practice.

[18] Choudhary, J. (2012). Survey of Different Biometrics Techniques, 2.

[19] Ganguly, S., & Moulick, S. (2012). A Review On Different Biometric Techniques. International Journal of Engineering Research & Technology, 1(5), 1–7.

[20] Naït-Ali, A., & Fournier, R. (2012). Signal and Image Processing for Biometrics. ISTE Ltd & Wiley.

[21] Li, H., Toh, K.-A., & Li, L. (2011). Advanced topics in biometrics. Advanced Topics in Biometrics. World Scientific.

Verificación y validación de representaciones visuales y sus interacciones

Martín Larrea, Matías Selzer, Dana K. Urribarri, M. Luján Ganuza, Silvia M. Castro
Laboratorio de I+D en Visualización y Computación Gráfica
Departamento de Ciencias e Ingeniería de la Computación, Universidad Nacional del Sur (DCIC-UNS)
Instituto de Ciencias e Ingeniería de la Computación (UNS-CONICET)
Laboratorio de I+D en Visualización y Computación Gráfica,
(UNS-CIC Prov. de Buenos Aires)
{mll, matias.selzer, dku, mlg, smc}@cs.uns.edu.ar

RESUMEN

Cada día más decisiones son tomadas en base al análisis visual mediante el uso de representaciones visuales de grandes conjuntos de datos. No sólo hay un aumento cuantitativo sino también cualitativo, las decisiones son más críticas y con mayor impacto en la sociedad, el medio ambiente y los individuos. En este contexto resulta fundamental desarrollar nuevas y mejores metodologías y herramientas que permitan asegurarnos el correcto funcionamiento de las representaciones visuales y sus interacciones.

Palabras Clave: Verificación y Validación, Visualización de Datos, Análisis Visual de Datos

CONTEXTO

Este trabajo se lleva a cabo en el Laboratorio de Investigación y Desarrollo en Visualización y Computación Gráfica (VyGLab, UNS-CIC Prov. de Buenos Aires) 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 y alumnos de grado.

1. INTRODUCCIÓN

La toma de decisiones se ha transformado en una herramienta clave en toda organización, pasando de ser un proceso basado en la experiencia y la intuición a uno cada vez más establecido en el análisis de datos. En este contexto, en el último tiempo se ha comenzado a hablar de Big Data; es el concepto que hace referencia a conjuntos de datos tan grandes y complejos donde los sistemas informáticos tradicionales de procesamiento de datos no pueden tratarlos adecuadamente. Por ende, los procedimientos usados para tomar decisiones en base a estos datos son más sofisticados y requieren software especializado. Un tipo de software especializado para Big Data, es el de representaciones visuales para grandes conjuntos de datos. Las representaciones visuales son de suma importancia en la actualidad ya que permiten la exploración efectiva de un conjunto de datos y facilitar la tarea de identificar patrones y extraer conclusiones. Los sistemas de representaciones visuales no sólo ofrecen una vista de los datos, sino también proveen interacciones mediante las cuales el usuario puede manifestar sus necesidades y a partir de las cuales obtener la perspectiva deseada del conjunto de datos. Cada día más decisiones son tomadas en base al análisis visual mediante el uso de representaciones visuales de grandes conjuntos de datos. No sólo hay un aumento cuantitativo sino también cualitativo, las decisiones son más críticas y con mayor impacto en la sociedad, el medio ambiente y los individuos. En este contexto resulta fundamental desarrollar nuevas y mejores metodologías y herramientas que permitan asegurarnos el correcto funcionamiento de las

representaciones visuales y sus interacciones ([1]).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Basándonos en el estado del arte de las representaciones visuales de grandes conjuntos de datos ([2]), y teniendo en cuenta sus aplicaciones ([3]), los estados intermedios de los datos ([4]) y las posibles interacciones ([5]), es posible definir e implementar metodologías y herramientas que permitan detectar errores y/o problemas en los softwares de representaciones visuales de grandes conjuntos de datos. La principal línea de investigación y desarrollo consiste en el estudio y desarrollo de metodologías para la validación y verificación de representaciones visuales de grandes conjuntos de datos. De esta línea principal se desprenden sublíneas relativas al diseño y desarrollo de herramientas que asistan en el desarrollo de representaciones visuales de grandes conjuntos de datos para asegurar la calidad del producto final. Estas herramientas buscarán eliminar o reducir la incidencia de errores en los sistemas de representación visual. Esta línea de investigación expande nuestro trabajo sobre el área de Visualización ([6, 7, 8, 9, 10]).

3. RESULTADOS OBTENIDOS Y ESPERADOS

Sobre los ejes presentados se han obtenido resultados parciales. Desde el punto de vista de la teoría de base se ha comenzado a realizar un relevamiento de las diferentes metodologías disponibles en el ámbito de la Verificación y Validación del Software que sean aplicables a un dominio tan particular como el de la Visualización de Información. En lo referido al desarrollo de nuevas metodologías y herramientas se ha comenzado a publicar los resultados obtenidos hasta el momento. Tal es el caso de "Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions" ([11]) y "White-Box Testing Framework for Object-Oriented Programming based on Message Sequence Specification" ([12]), este último seleccionado como mejor trabajo para la "Communications in Computer and Information Science" de Springer. En el corriente año se pretende avanzar sobre una herramienta específica para el testeo de visualizaciones y la publicación de nuevas metodologías.

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 evaluación

Tesis Desarrollada: "TÉCNICAS DE TESTING ORIENTADAS A OBJETOS BASADAS EN SECUENCIAMIENTO DE MÉTODOS", tesis de grado para la Licenciatura en Ciencias de la Computación. Alumno: Juan Ignacio Rodríguez Silva. Director: Martín Larrea, Co-Directora: Dana K. Urribarri, Colaborador: Matías Selzer

Tesis en Desarrollo: "GENERACIÓN DE DATOS DE PRUEBA PARA VERIFICACIÓN Y VALIDACIÓN DE SOFTWARE EN LA INDUSTRIA MEDIANTE ALGORITMOS GENÉTICOS", tesis de grado para la Ingeniería en Sistemas de Información. Alumna: Constanza Giorgetti. Director: Martín Larrea, Colaborador: Matías Selzer. Para esta alumna también se solicitó una Beca de Inicio en la Investigación de la Universidad Nacional del Sur. La misma se encuentra en evaluación.

Proyecto: PGI 2019 “VERIFICACIÓN Y VALIDACIÓN DE REPRESENTACIONES VISUALES Y SUS INTERACCIONES”. Director: Martín Larrea, Miembros: Matías Selzer, Dana K. Urribarri, M. Clara Casalini. Proyecto presentado para acreditación de incentivos, en evaluación.

5. BIBLIOGRAFÍA

[1] Kirby, Robert M., and Cláudio T. Silva. “The need for verifiable visualization”. Computer Graphics and Applications, IEEE 28.5, 78-83, 2008.

[2] Rees, D., and R. S. Laramee. "A Survey of Information Visualization Books." Computer Graphics Forum. 2019.

[3] Börner, Katy, Andreas Bueckle, and Michael Ginda. "Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments." Proceedings of the National Academy of Sciences 116.6 (2019): 1857-1864.

[4] Hagen, Hans. "DFU, Volume 2, Scientific Visualization: Interactions, Features, Metaphors, Complete Volume." Dagstuhl Follow-Ups. Vol. 2. Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2012.

[5] Ganuza, M. L. "Thesis Overview: Interactions in Visualization." Journal of Computer Science & Technology 18.2 (2018): 178-179.

[6] Castro, Silvia Mabel, et al. "Métricas, técnicas y semántica para la visualización de datos." XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).. 2018.

[7] Urribarri, Dana K., and Silvia M. Castro. "Prediction of data visibility in two-dimensional scatterplots." Information Visualization 16.2: 113-125. 2017

[8] Ganuza, María Luján, et al. "Visualización de Datos aplicada a resolver problemas en Ciencias Geológicas." XVI Workshop de Investigadores en Ciencias de la Computación. 2014.

[9] Larrea, Martín Leonardo, et al. "Ontologías y Semántica en el Proceso de Visualización." XVI Workshop de Investigadores en Ciencias de la Computación. 2014.

[10] Selzer, Matías Nicolás, et al. "Modelos de interacción y aplicaciones en realidad virtual mediante dispositivos móviles." XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires). 2017.

[11] Larrea, Martín Leonardo. "Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions." Journal of Computer Science & Technology 17. 2017.

[12] Silva, J.I.R., Larrea, M.: White-box testing framework for object-oriented programming based on message sequence specification. In: XXIV Congreso Argentino de Ciencias de la Computación (Tandil, 2018). 2018.

Visualización de publicaciones científicas empleando D3

José Federico Medrano[1], José Luis Alonso Berrocal[2], Carlos G. Figuerola[2]

jfmedrano@fi.unju.edu.ar, berrocal@usal.es, figue@usal.es

[1]VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

[2] REINA / Recuperación de Información Avanzada / Facultad de Traducción y Documentación

Universidad de Salamanca – España - C/ Francisco de Vitoria, 6-16

RESUMEN

Las bases de datos bibliográficas poseen una enorme cantidad de registros en forma de publicaciones científico-académicas, al realizar una búsqueda por autor, por ejemplo, ofrecen un conjunto de resultados, el cálculo de indicadores y métricas, y en algunos casos opciones de visualización de resultados. Este último aspecto, la visualización, ha sido un poco descuidado. Estas herramientas se limitan a presentar simples estadísticas o datos tabulados, o los conocidos grafos de co-autor o co-citas, dejando de lado algún otro tipo de visualización más compleja como la evolución temporal involucrando varias dimensiones más que la cantidad de citas recibidas por año. Una adecuada visualización de datos permitirá entender la realidad desde distintas ópticas aportando un mayor entendimiento y conocimiento a veces oculto en representaciones básicas y estáticas.

El desarrollo de este proyecto plantea encontrar la mejor forma, la más adecuada o proponer una nueva visualización para representar las relaciones existentes entre autores, publicaciones, citas recibidas, y cualquier otro tipo de información o relación que se considere relevante durante el estudio.

Palabras clave: *Data Visualization; Publicaciones científicas; DataVis; Visualización temporal*

CONTEXTO

La línea de investigación aquí presentada se encuentra enmarcada dentro del Proyecto Consolidado D/B029 denominado "*Aplicación de técnicas de Inteligencia Artificial para evaluar la producción científico-académica de investigadores de Universidades públicas del Noroeste Argentino*", aprobado y financiado por la Secretaría de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy. Esta etapa del proyecto es llevado a cabo en conjunto por dos grupos de investigación. En primer lugar liderado por el grupo de investigación VRAIn de la Facultad de Ingeniería de la Universidad Nacional de Jujuy, y en segundo lugar como colaboradores, el grupo REINA de la Facultad de Traducción y Documentación de la Universidad de Salamanca

1. INTRODUCCIÓN

En las décadas anteriores era muy común que los datos se presentaran textualmente o mediante gráficos estáticos, en estos casos la información representada estaba limitada a cantidades

pequeñas, pero en los últimos años estos tipos de representaciones se han tornado poco útiles cuando se trata de conjuntos de datos que contienen millones de elementos de datos (Keim, 2002). Los sistemas actuales almacenan grandes cantidades de datos y al no tener la posibilidad de explorarlos adecuadamente, los datos se vuelven inútiles y las bases de datos se convierten en meros depósitos.

La aparición de interfaces gráficas ha permitido una interacción directa con la información visualizada, dando lugar a más de una década de investigación en Visualización de Información (InfoVis) (Heer, Card, & Landay, 2005). InfoVis busca aumentar el conocimiento humano mediante el aprovechamiento de las capacidades visuales humanas para dar sentido a la información abstracta (Card, Mackinlay, & Shneiderman, 1999), proporcionando los medios por los cuales los seres humanos mediante sus capacidades perceptivas, pueden lidiar con el constante aumento de la cantidad de datos disponibles.

El objetivo de InfoVis es profundizar en los datos o conceptos ocultos. A menudo la información se oculta simplemente por la enorme cantidad de datos disponibles. De este modo, la Visualización de Información también puede ser vista como un convertidor entre los datos subyacentes y la percepción humana de la misma (Prinz, 2006). El representar grandes cantidades de información mediante abstracciones no es una tarea fácil ya que el usuario no tiene ninguna idea preconcebida de cómo estos datos pueden ser representados.

La idea básica de la exploración visual de los datos es la de presentar los datos en alguna forma visual, permitiendo que los humanos puedan obtener conocimiento, sacar conclusiones, e interactuar directamente con los mismos. Con este tipo de representaciones basadas en grandes cantidades de datos, los usuarios pueden detectar patrones o comportamientos que se deseaban evaluar, como así también descubrir comportamientos y relaciones entre los datos desconocidos hasta el momento

La interacción es especialmente importante en InfoVis, ya sea para la exploración, análisis y/o presentación de los datos (Kosara, Hauser, & Gresh, 2003). La interacción permite al usuario implícitamente formar modelos mentales de las correlaciones y las relaciones entre los datos, a través del reconocimiento de patrones. El uso de los ordenadores permite ir un paso más allá de simples representaciones del mundo real, permitiendo hacer agrupaciones o asociaciones impensables, o distorsiones sobre dichas representaciones proporcionando un mayor nivel de abstracción, aspecto fundamental de las técnicas de InfoVis.

La Visualización de Información cumple un papel relevante al realizar cualquier estudio del estado de la producción científico-académica de un investigador, y más aún al considerar el recuento de citas de dichas publicaciones. Poder plasmar mediante una representación, la información recolectada se vuelve una tarea compleja no solo según aumente la cantidad de información, sino también según se incremente el número de dimensiones objeto de estudio. Las bases de datos bibliográficas hoy en día ofrecen un pequeño número de visualizaciones relacionadas con esta temática, se limitan a un conjunto de gráficos estáticos como la evolución de algunos indicadores calculados o el recuento o clasificación de las publicaciones de un autor en particular (Medrano, 2017).

El diseño de visualizaciones para la exploración de datos temporales requiere varias opciones basadas en aspectos de tiempo y representación visual. La elección del diseño o técnica debe adecuarse al problema en cuestión y a los valores que se desean informar, la misma magnitud puede representarse de múltiples formas y el

conocimiento que se pueda obtener de ella no tendrá el mismo impacto visual aplicando un diseño u otro. Como indica (Henkin, Dykes, & Slingsby, 2016) los aspectos a tener en cuenta deberían basarse en el diseño, la forma y el tamaño de las marcas visuales. En (Bach, Dragicevic, Archambault, Hurter, & Carpendale, 2016; 2014) se presenta una taxonomía completa de las distintas técnicas y modelos utilizados para representar datos temporales, tanto en 2 dimensiones (2D) como en 3 dimensiones (3D).

El *Scatter plot* también llamado *scatter graph*, *scatter chart*, *scattergram*, *scatter diagram* o diagrama de dispersión, es un tipo de diagrama matemático que utiliza coordenadas cartesianas para graficar puntos que muestran la relación entre dos variables de un conjunto de datos. Los puntos pueden ser coloreados para indicar los valores de una variable adicional. Al ser puntos dentro de un eje de coordenadas, el valor de cada punto está dado por la posición que ocupa, es decir, el valor de una variable según la posición en el eje horizontal y el valor de la otra variable según la posición en el eje vertical. A menudo se utiliza este tipo de diagramas para identificar asociaciones potenciales entre dos variables, en las que se puede considerar una variable explicativa (como años de educación) y otra puede considerarse una variable de respuesta (como el ingreso anual) (Lacey, 2017).

Los conjuntos de datos que involucran más de dos magnitudes resultan difíciles de visualizar en un espacio de 2D, en algunos casos visualizaciones en 3D pueden aportar ciertas mejoras o nuevas posibilidades, sin embargo al aumentar la cantidad de magnitudes o dimensiones del problema resulta necesario recurrir a visualizaciones un tanto más complejas que una representación lineal en ejes de coordenadas. Para este trabajo se diseñó una visualización en el espacio 2D recurriendo a ciertos elementos para poder incluir todas las magnitudes que debían ser representadas para capturar, en una sola representación, el entendimiento global del estado actual de los resultados de la investigación de un científico.

Para esto se desarrolló un primer prototipo de visualización utilizando la librería D3.js (Data Driven Documents) (Teller, 2013; Zhu, 2013).

D3 (creada por Mike Bostock[1]) es una elegante pieza de software que facilita la generación y manipulación de documentos web con datos. Lo realiza mediante (Murray, 2013):

- Carga de datos en la memoria del navegador.
- Vincular datos a elementos dentro del documento, creando nuevos elementos según sea necesario.
- Transformar esos elementos mediante la interpretación de los datos enlazados de cada elemento y establecer sus propiedades visuales en consecuencia.
- Transición de elementos entre estados en respuesta a la entrada del usuario.

D3.js fue creado para llenar una necesidad apremiante de una sofisticada visualización de datos accesible desde la web. La visualización de datos ya no se refiere a los gráficos de torta y gráficos de líneas. Ahora significa mapas y diagramas interactivos y otras herramientas y contenidos integrados en noticias, cuadros de mando de datos, informes y todo lo que se ve en la web.

La visualización desarrollada se basó en un ejemplo sencillo y estático de *Scatter plot* disponible en el blog de su creador (Bostock, 2019), ver Figura 1. Por su parte, en la Figura 2

[1] https://bost.ocks.org/mike/

se observa el primer prototipo de la visualización desarrollada.

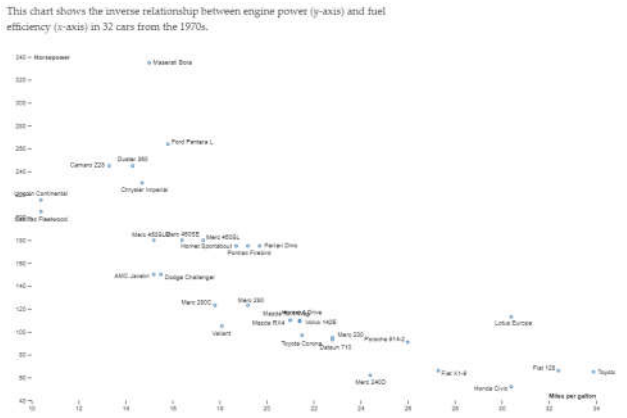


Figura 1: Visualización de Scatterplot provista por Mike Bostock

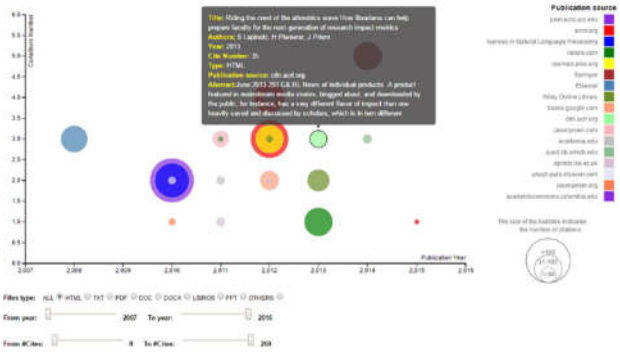


Figura 2: Visualización de Scatterplot adaptado

En el eje vertical se encuentra la cantidad de coautores de las publicaciones, en el eje horizontal el año de publicación, cada publicación se representa por una burbuja donde el tamaño de las mismas indica la cantidad de citas recibidas, y por último el color de las burbujas indica el nombre de la revista donde fue publicado el registro bibliográfico. Este prototipo ofrece un conjunto de opciones de interacción entre las que se incluye la posibilidad de filtrar por el origen de la publicación (mostrando u ocultando las publicaciones de dicho origen), filtrar por el tipo de archivo de la publicación eligiendo de la lista disponible, filtrar los registros por año de publicación, filtrar las burbujas de acuerdo a la cantidad de citas recibidas. Además permite visualizar el detalle de la publicación al pasar el mouse por encima de una burbuja y acceder al recurso publicado haciendo doble-clic sobre la burbuja.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de trabajo principal de este proyecto de investigación es el estudio, diseño e implementación de técnicas de Visualización de Información aplicables a representar el conjunto de publicaciones científico-académicas de un investigador a partir de la recolección de registros bibliográficos almacenados en bases de datos académicas tanto comerciales (Scopus, Web of Science) como de libre acceso (Google Scholar, Microsoft Academic), con el objeto de mejorar los procesos de comunicación de resultados de los análisis bibliométricos que se realizan sobre estas bases de datos. A partir de la generación de visualizaciones de dichos resultados, se logrará una mejor comprensión de los datos.

3. RESULTADOS OBTENIDOS/ESPERADOS

Con este proyecto se espera encontrar un modo novedoso que satisfaga la enorme necesidad de contar con visualizaciones de información de publicaciones científicas que permitan brindar una óptica diferente a resultados tabulares o gráficos sencillos y estáticos.

Particularmente se espera lograr:

- Mejorar los procesos de comunicación de resultados de los análisis bibliométricos que se realizan sobre bases de datos bibliográficas.
- Mejorar la comprensión de los datos a partir de la generación de visualizaciones de dichos resultados.
- Diseñar una herramienta web de libre acceso que permita generar visualizaciones a partir de un dataset específico.
- Diseñar una herramienta interactiva, agradable e intuitiva que ofrezca un

conjunto de opciones para aumentar el entendimiento sobre la información presentada.

4. FORMACIÓN DE RECURSOS HUMANOS

Este proyecto brinda un marco para que docentes y estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico.

El área de visualización de información es incluida y desarrollada en cada proyecto que los equipos de investigación llevan a cabo, además de presentar un campo atractivo y novedoso a los alumnos que forman y formarán parte de los distintos proyectos.

5. BIBLIOGRAFÍA

Bach, B., Dragicevic, P., Archambault, D., Hurter, C., & Carpendale, S. (2014). A review of temporal data visualizations based on spacetime cube operations. *Eurographics Conference on Visualization.*

Bach, B., Dragicevic, P., Archambault, D., Hurter, C., & Carpendale, S. (2016). A descriptive framework for temporal data visualizations based on generalized space-time cubes. *Computer Graphics Forum.*

Bostock, M. (2019). *Observable*. Obtenido de https://observablehq.com/@d3/scatterplot

Card, S., Mackinlay, J., & Shneiderman, B. (1999). *Readings in Information Visualization: Using Vision to Think.* San Francisco: Morgan-Kaufmann.

Heer, J., Card, S. K., & Landay, J. (2005). Prefuse: A toolkit for interactive information visualization. *ACM Human Factors in Computing Systems (CHI)*, 421-430.

Henkin, R., Dykes, J., & Slingsby, A. (2016). Characterizing representation of temporal data visualization. *Poster presented at the VIS 2016*.

Keim, D. A. (2002). Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics, 7*(1).

Kosara, R., Hauser, H., & Gresh, D. (2003). An interaction view on information visualization. *EUROGRAPHICS 2003 State-of-the-Art Re-ports*, (págs. 123-137).

Lacey, M. (2017). *Statistical topics*. Obtenido de http://www.stat.yale.edu/Courses/1997-98/101/scatter.htm

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.

Murray, S. (2013). *Interactive Data Visualization for the Web.* O'Reilly.

Prinz, W. (2006). The Graph Visualization System (GVS): A Flexible Java Framework for Graph Drawing. *Master's thesis, Graz University of Technology*.

Teller, S. (2013). *Data Visualization with D3.js.* Packt Publishing.

Ware, C. (2004). *Information Visualization - Perception for Design.* Morgan-Kaufmann.

Ware, C. (2008). *Visual Thinking for Design.* Morgan Kaufman/Elsevier.

Zhu, N. Q. (2013). *Data Visualization with D3.js Cookbook.* Packt Publishing.

Ingeniería

de Software

ABORDAJE INFORMÁTICO EN EL DISEÑO E IMPLEMENTACIÓN DE SISTEMAS DE IOT, GARANTIZANDO SEGURIDAD Y ESCALABILIDAD

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.flores@ieee.org, { mberon, driesco }@unsl.edu.ar
pedrorangelhenriques@gmail.com

RESUMEN

En la actualidad, los avances en el dominio del Internet of Things están transformando las formas en que las personas interactúan con el entorno y están dotando de inteligencia al mismo. Estos progresos buscan mejorar la calidad de vida de las personas, optimizar los procesos industriales e incrementar el cuidado del medio ambiente y de sus recursos. Grandes empresas han implementado soluciones de IoT en sus líneas de producción con el fin de incrementar su competitividad de cara a la industria 4.0. Sin embargo, el correcto desarrollo de un sistema de IoT podría ser un desafío si no se toman ciertos recaudos. Uno de los problemas que se presentan, desde la planificación del sistema, es el diseño de una arquitectura que favorezca la escalabilidad en cuanto a cantidad de clientes del sistema, ubicación geográfica de los mismos y, en la variedad de dispositivos y aplicaciones conectadas. Por otro lado, un sistema de IoT debe poder garantizar eficiencia y seguridad en comunicaciones persona-máquina y máquina-máquina.

En este artículo, se presenta una línea de investigación que aborda el diseño y construcción de *sistemas de IoT Escalables y Seguros,* buscando superar las dificultades mencionadas.

Palabras clave: *Sistema de IoT, Dispositivo, Cosa, Escalabilidad, Seguridad, Privacidad, Arquitectura, Internet.*

CONTEXTO

La presente línea de investigación se en marca 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 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" (i.e. Internet de las Cosas). El mismo corresponde a un dominio de aplicación que integra diferentes campos tecnológicos y sociales. De acuerdo con IEEE et. al. [1], 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.

Concordando con IEEE et. al. en [1], existen

varios puntos importantes a destacar en el Software de los sistemas de IoT:

- Los sistemas operativos usados en dispositivos IoT deben estar diseñados para ejecutarse de forma eficiente (ya que trabajan en componentes de pequeña escala), sin dejar de proporcionar funcionalidades básicas que den soporte a los objetivos y propósitos de las aplicaciones ejecutadas en ellos [2, 3, 4].

- 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 favorezcan la reutilización de componentes y una adecuada gestión de los datos a almacenar/procesar [2, 5, 6, 7].

- Los sistemas de IoT, potencialmente pueden crecer y llegar a componerse por millones de dispositivos diferentes, cada uno ubicado en lugares remotos del planeta y con usuarios ubicados también remotamente. En este contexto, la autogestión y auto-optimización de cada dispositivo y/o subsistema individual podría convertirse en la norma [2, 3, 8].

- La seguridad y privacidad deben ser garantizadas en cada proceso de los sistemas de IoT, ya que los mismos producen y manejan información reservada [7, 9, 10, 11].

- 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. Tal interconexión 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, 8, 12, 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.

Más allá de las características particulares de cada uno, los sistemas de IoT pueden ser utilizados en los siguientes ámbitos:

- Sistemas de domótica.

- Extracción y análisis automáticos o semi automáticos de datos en líneas de producción industrial, que permitan optimizar los procesos de negocio [14].

- Monitoreo del estado físico de personas o animales, o extensión de sus capacidades físicas, a través de dispositivos que puedan ser vestidos (i.e. wearables).

- Monitoreo de cultivos y riego inteligente.

- Monitoreo medio-ambiental.

Los sistemas de IoT no necesariamente requieren de interacción humana en su funcionamiento. ETSI [12] 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 de supervisión de 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.

Por 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 para permitir modificaciones estructurales del sistema, de una forma comprensible y segura. También, debe permitir la obtención de múltiples vistas que representen 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. [14], llevan a que un 60% de las iniciativas IoT se estanquen en la fase de prueba de concepto y, del 40% restante, sólo el 26% sean consideradas un completo éxito.

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. De hecho, Kranz et al. estima que, en el año 2020, se alcanzará una cifra de 50.000 millones de objetos conectados en todo el mundo [14].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la actualidad, existe una gran variedad de sistemas operativos destinados a dispositivos IoT (e.g. TinyOS, Raspbian, RIOT, Ubuntu, Windows 10 IoT Core) [1, 4]. 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 et. al. [1], frente al crecimiento en la cantidad de *cosas conectadas a internet*, 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.

Por otro lado, mientras que IoT promete una mejor vida a través de dispositivos conectados y de la información que ellos generen, también marca el comienzo de una nueva era en cuanto a la privacidad y la seguridad. Según OWASP [15], uno de los nuevos riesgos de seguridad asociados al 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:

- Investigar en profundidad 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 desarrolló F-IoT Core, una

herramienta para el desarrollo de soluciones IoT escalables y flexibles [16]. La misma posee una arquitectura que permite implementar soluciones IoT de cualquier escala, con la suficiente flexibilidad para modificarlos en versiones posteriores, sin la necesidad de reescribir a todos sus módulos funcionales. Por otro lado, la herramienta brinda la posibilidad de elegir entre dos protocolos estándar del IoT (MQTT y HTTP) para realizar las comunicaciones, con el fin de tener un mayor alcance y de brindar la posibilidad de realizar conexiones persistentes y no persistentes, dependiendo del protocolo que se use.

F-IoT Core fue desarrollada en Java y está destinada a ser instalada en un Servidor Web, con el fin de brindar soporte a las comunicaciones entre dispositivos IoT y usuarios remotos. Se eligió Java como lenguaje de programación, ya que, en la actualidad, la máquina virtual de Java es un estándar que puede ser ejecutado en casi cualquier sistema operativo de forma eficiente y segura.

A futuro se poseen los siguientes objetivos:

- Desarrollar un frontend utilizando Angular como tecnología Web [17], para facilitar la configuración de sistemas de IoT a través de F-IoT Core, ya que en la actualidad esto se realiza a través de una API REST. Además, este frontend debe permitir la visualización del estado de los sistemas de IoT, a través de diferentes representaciones/vistas.

- Realizar una investigación profunda acerca de las amenazas de seguridad que afectan a los sistemas de IoT, y crear una capa de seguridad que abarque todas las comunicaciones realizadas en F-IoT Core.

- 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.

- Analizar y comparar las diferentes tecnologías IoT disponibles para ejecución en la Nube, a través de servicios provistos por terceros (e.g. Microsoft, Google, Amazon, IBM, Oracle).

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 mencionados en la Sección *Contexto*.

5. BIBLIOGRAFÍA

[1] IEEE, «Towards a Definition of the Internet of Things (IoT),» 27 05 2015. [En línea]. Available: https://iot.ieee.org/images/files/pdf/IEEE_IoT_Towards_Definition_Internet_of_Things_Revision1_27MAY15.pdf. [Último acceso: 06 06 2018].

[2] IETF, «The Internet of Things - Concept and Problem Statement,» 2010. [En línea]. Available: https://www.ietf.org/archive/id/draft-lee-iot-problem-statement-05.txt.

[3] CERP-IoT, «Visions and Challenges for Realising the Internet of Things,» 2010. [En línea]. Available: http://www.internet-of-things-research.eu/pdf/IoT_Clusterbook_March_2010.pdf.

[4] P. G. M. P. T. S. Pérez, «Operating Systems for IoT Devices: A Critical Survey,» *2015 IEEE Region 10 Symposium,* 05 2015.

[5] 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.

[6] H. Ren, H. Li, Y. Dai, K. Yang y X. Lin, «Querying in Internet of Things with Privacy Preserving: Challenges, Solutions and Opportunities,» 13 Marzo 2018. [En línea]. Available: http://ieeexplore.ieee.org/document/8315210/.

[7] T. Choudhury, A. Gupta, S. Pradhan, P. Kumar y Y. S. Rathore, «Privacy and Security of Cloud-Based Internet of Things (IoT),» 28 Octubre 2017. [En línea]. Available: http://ieeexplore.ieee.org/document/8307328/.

[8] N. B. A. C. L. M. Z. A. Zanella, «Internet of Things for Smart Cities,» *IEEE Internet of Things Journal,* 02 2014.

[9] 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 Febrero 2018. [En línea]. Available: http://ieeexplore.ieee.org/document/8279412/.

[10] M. El-hajj, M. Chamoun, A. Fadlallah y A. Serhrouchni, «Taxonomy of authentication techniques in Internet of Things (IoT),» Diciembre 2017. [En línea]. Available: http://ieeexplore.ieee.org/document/8305419/.

[11] J. Singh, T. Pasquier, J. Bacon, H. Ko y D. Eyers, «Twenty Security Considerations for Cloud-Supported Internet of Things,» *IEEE Internet of Things Journal ,* 2015.

[12] ETSI, «ETSI Technical Specification, "Machine-to-Machine Communications (M2M); M2M Service Requirements." Technical Specification.,» 08 2010. [En línea]. Available: www.etsi.org/deliver/etsi_ts/102600_102699/102689/01.01.01_60/ts_102689v010101p.pdf. [Último acceso: 04 09 2018].

[13] D. Zeng, S. Guo y Z. Cheng, «The Web of Things: A Survey (Invited Paper),» 2011. [En línea]. Available: http://www.jocm.us/index.php?m=content&c=index&a=show&catid=50&id=134.

[14] M. Kranz, Internet Of Things. Construye nuevos modelos de negocio., LID Editorial, 2017.

[15] OWASP, «IoT Attack Surface Areas Project,» [En línea]. Available: https://www.owasp.org/index.php/OWASP_Internet_of_Things_Project#tab=IoT_Attack_Surface_Areas.

[16] S. Flores, M. Berón, D. Riesco, P. Henriques y M. Bustos, «F-IoT Core, una Herramienta para el Desarrollo de Soluciones IoT Escalables y Flexibles,» de *CoNaIISI*, Mar del Plata, 2018.

[17] Google, «Angular,» [En línea]. Available: https://angular.io/.

Accesibilidad web visual. Formación de RRHH para fomentar su desarrollo en la región NEA

Sonia I. Mariño, Pedro L. Alfonzo, Verónica K. Pagnoni
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, vero_pagnoni@hotmail.com

Resumen

Se exponen proyectos relacionados con la formación de RRHH de grado y posgrado en torno a la Accesibilidad web visual, un aspecto de la calidad del software. Los conocimientos adquiridos a través de la formación de estos sujetos que se desarrollan en la industria del software y los productos tecnológicos que se deriven de los saberes que se construyan, favorecerá a la mejora de los productos que consumen los e-ciudadanos en el NEA.

Palabras clave: Accesibilidad Web, accesibilidad visual, métodos y herramientas, formación de recursos humanos, transferencia de conocimientos.

Contexto

Esta propuesta se desarrolla en el marco del proyecto de I+D "Sistemas de Información y TIC: métodos y herramientas" acreditado por la Secretaria General de Ciencia y Técnica (SGCyT – UNNE) y ejecutado por un equipo de docentes – investigadores de la FACENA – UNNE. Este equipo de trabajo originado en 2005 diseña y ejecuta distintas actividades relacionadas con el estudio y construcción de soluciones con miras a la transferencia de conocimientos. Particularmente en esta ponencia se describe el trabajo realizado por los integrantes en torno a indagar y aplicar métodos y herramientas para evaluar la accesibilidad web visual, según estándares internacionales como un aspecto de la calidad de la Ingeniería del Software

Introducción

La Accesibilidad Web (AW) 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 [1, 2].

En [3] definen la Accesibilidad Web Accesibilidad Web o la posibilidad de que un producto o servicio web pueda ser accedido y usado por el mayor número posible de personas, indiferentemente de las limitaciones propias del individuo o de las derivadas del contexto de uso. De este modo, la AW no es una característica deseable sino un requisito necesario, debido al aumento creciente de la participación activa de usuarios con capacidades diferentes [4]. Es así como se identifican diferentes discapacidades que pueden impedir el acceso a la información de los contenidos web, como ser: i) Deficiencias visuales; ii) Deficiencias auditivas; ii) Deficiencias motrices; iv) Deficiencias cognitivas y de lenguaje. Además, cada una de las categorías de discapacidad mencionadas, requerirá de

atención específica cuando se desarrollen productos web [3]. En este sentido, respecto a las deficiencias visuales, se deben considerar la ceguera, la visión reducida y los problemas en visualización de color. El estudio y tratamiento de la AW contemplando esta discapacidad cobra trascendencia dado que mundialmente se estima que 1300 millones de personas poseen alguna forma de deficiencia visual [5].

En el marco del proyecto de "Sistemas de Información y TIC: 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.

Es Responsabilidad Social de los actores de la Universidad colaborar con la consolidación de la sociedad de la Información [6, 7, 8]. Así involucrar en el grado y posgrado a los futuros profesionales, se considera de relevancia para la Industria del Software.

Desde la perspectiva de la Ingeniería del Software (IS) es notable 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 [8].

Para el equipo es fundamental el estudio teórico y la definición de procesos orientados a aplicar la AW a lo largo de desarrollos tecnológicos [9, 10, 11] en particular aquellos que abordan la Accesibilidad web visual dado el elevado número de personas a las que comprende.

Cabe aclarar que la iniciativa de estudio de AW, desarrollada por este equipo de trabajo de la UNNE, coincide con el tratado en otras universidades [12, 13, 14].

Líneas de Investigación, Desarrollo e Innovación

Este equipo aborda las siguientes líneas de investigación:

- Revisión de la literatura relacionada con propuestas y estudios de casos relativos a la AW y en particular a la Accesibilidad web visual.
- Relevamiento y revisión de normas y estándares internacionales referentes a la AW.
- Selección de herramientas TIC para la mejora de la Accesibilidad web visual.
- Determinación de equipos de I+D existentes que abordan la temática relacionada con este proyecto para proponer vínculos de cooperación-colaboración con ellos.

Resultados y Objetivos

Los objetivos establecidos en planes de formación de recursos humanos de grado y posgrado se definen como:

A. Determinar y seleccionar literatura y antecedentes en torno a la AW y en particular a la Accesibilidad web visual.
B. Relevar normas y estándares internacionales referentes a la AW.
C. Seleccionar herramientas TIC para la mejora de la Accesibilidad web visual en entornos web.
D. Identificar equipos de trabajos que aborden la temática.

Se lograron avances en los siguientes tópicos:

A. Determinar en la literatura y antecedentes en torno a la AW y en particular a la accesibilidad web visual:

Se realiza una extensa revisión y selección de la literatura referida a la AW y en particular a la visual. Las fuentes de información recopilada y sistematizada dan cuenta de avances y posibles áreas de trabajo.

B. Identificar normas y estándares relacionados con la AW:

- Sistema de normalización internacional para productos de áreas diversas definido por la Organización Internacional de Estandarización (ISO) aquellas relacionadas con la AW [15]
- Las relacionadas al Estándar WC3 para desarrollo web y móvil [16].
- Las pautas y recomendaciones de las Normas UNE de AENOR [17].

C. Seleccionar herramientas TIC para la mejora de la Accesibilidad web visual en entornos web:

En referencia a esta objetivo se avanzó con la elección y estudio de las herramientas TI orientadas a [18, 19, 20, 21]:

- Revisión automática de la AW.
- Revisión manual de la AW.
- Validación de contenido.
- Validación de código.

D. Identificar equipos de trabajos que aborden la temática:

En referencia a este objetivo se avanza implícitamente a partir de las publicaciones relevadas.

Formación de Recursos Humanos

En la región NEA, desde el ámbito académico se tiene un fuerte compromiso de aportar al desarrollo de la Industria del Software.

Asimismo, esta línea de I+D+i se corresponde a un proyecto en el cual la formación de recursos humanos de grado y de posgrado es fundamental. El equipo de trabajo formado por docentes – investigadores integró a un alumno de proyecto de finalización de la carrera Licenciatura en Sistemas de Información quien defendió su trabajo en el año 2018. Además, se fortalece esta línea de I+D+i a través de dos planes de tesis de posgrado. De esta manera se contribuye a fortalecer la transferencia de conocimientos desde el equipo original dado que estos sujetos serán replicadores del conocimiento adquirido y experimentado.

Cabe recordar algunos trabajos que evidencian la formación de sujetos en la temática [22, 23, 24, 25].

Derivado de esta formación y como línea futura se prevé asesorar con la finalidad de aportar, desde la producción de software en la conformación de una sociedad inclusiva centrada en los sujetos consumidores de la tecnología.

Referencias

[1] Consorcio World Wide Web (W3C). Disponible en: http://www.w3c.es/

[2] Oficina Española. (2008). Word Wide Web - Guía Breve de Accesibilidad Web. Disponible en: http://www.w3c.es/divulgacion/guiasbreves/accesibilidad

[3] Hassan Montero, Y. y Martín Fernández, F. J. (2003). “Qué es la Accesibilidad Web”. No Solo Usabilidad, 2.

[4] Miranda, G., Martin, A., Mazalu, R., Gaetán, G., & Saldaño, V. (2013). *Agentes Inteligentes para propiciar la Accesibilidad Web.* Neuquén, Argentina: Universidad Nacional de la Patagonia Austral-Universidad Nacional del Comahue.

[5] Organización Mundial de la Salud. (octubre de 2018). *Ceguera y*

discapacidad visual. Disponible en: https://www.who.int/es/news-room/fact-sheets/detail/blindness-and-visual-impairment

[6] Toca Torres, C. E. (2017). "Aportes a la responsabilidad social". Revista Mexicana de Ciencias Políticas y Sociales. LXII, 230, pp. 393-408.

[7] Larrán, J. M., Andrades Peña, F. J. (2015). "Análisis de la responsabilidad social universitaria desde diferentes enfoques teóricos". Revista Iberoamericana de Educación Superior, 6(15), pp. 91-107.

[8] Hernández-Arteaga, R. I., Alvarado-Pérez, J. C. y Luna, J. A. (2015). Responsabilidad social en la relación universidad-empresa-Estado. Educación y Educadores. 18(1), pp. 95-110.

[9] Mariño, S. I., Godoy, M. V. y Alfonzo, P. L. (2017). "Avances en torno a la formación en accesibilidad web". XIX Workshop de Investigadores en Ciencias de la Computación.

[10] Mariño, S. I.; Alfonzo, P. L. (2017). Evaluación de la accesibilidad web. Una mirada para asegurar la formación en la temática. Campus Virtuales, 6(2), pp. 21-30.

[11] Mariño S. I., Alfonzo P.L. (2017). Inclusión Social: TIC y Accesibilidad Web. Primer Congreso de Educación y Tecnologías del Mercosur, Corrientes, Argentina.

[12] Barrios, T., Marín, M. y Torrente, N. (2017). "El Uso de la Tecnología para la Inclusión de los Disminuidos Visuales en las Aulas". XIX Workshop de Investigadores en Ciencias de la Computación.

[13] Amatriain, H., Martins, S., Bianco, S., Ribeiro, F. y Pérez, N. (2018). "Herramientas de interpretación gráfica para codificación de diagramas de modelado de sistemas interpretables por un disminuido visual". XX Workshop de Investigadores en Ciencias de la Computación.

[14] Lafuente, G. J., Ballesteros C. y Filippi, J. L. (2018). "Desarrollo de una plataforma de capacitación a distancia utilizable y abierta para personas con discapacidad visual". XX Workshop de Investigadores en Ciencias de la Computación.

[15] ISO, www.iso.org

[16] W3C. (2015). Accesibilidad móvil: cómo se aplican WCAG 2.0 y otras pautas de W3C / WAI al móvil. Obtenido de https://www.w3.org/TR/mobile-accessibility-mapping/

[17] AENOR. (2012). UNE 139803: Requisitos de accesibilidad para contenidos en la Web. Madrid, España: AENOR.

[18] Check my colours (2015). Checkmycolours. Obtenido de: http://www.checkmycolours.com/

[19] Access Keys. (s.f.). AccessColor - Herramienta Online para el contraste de color. Obtenido de http://www.accesskeys.org/tools/color-contrast.html

[20] SEDIC. (s.f.). Accesibilidad Web: Técnicas y herramientas para evaluar la accesibilidad web. Obtenido de http://www.sedic.es/autoformacion/accesibilidad/9-tecnicas-herramientas.html

[21] iamcal.com. (s.f.). Color Vision. Obtenido de https://www.iamcal.com/misc/colors/

[22] Blanco Vargas, V. Tema: "Sistema web de gestión de turnos para una clínica de ojos", Dir. Alfonzo, P.

L. Proyecto Final de Carrera, FaCENA, UNNE, Inédito, 2017.

[23] Mariño S. I., Alfonzo P.L., Galain, C., Maidana, J., Alderete, R. (2018). "Accesibilidad web, aportando a la inclusión". XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018). Argentina.

[24] Mariño S. I., Alfonzo P.L., Gomez Codutti, A. E. (2018). "An empirical WCAG 2.0 guidelines evaluation applied to journal management software". International Journal of Recent Engineering Research and Development (IJRERD). 3(11).

[25] Pagnoni, V. K. (2018) "Marco de trabajo de evaluación de la accesibilidad web para entornos virtuales de enseñanza-aprendizaje", Tesis de Ingeniería del Software, Dir. Mariño, S. I., Rossi, G. Trabajo de Maestria en Ingenieria del Software, UNLP, (en realización).

Algoritmia para Computadores Cuánticos

Samira Abdel Masih[1], Enrique Cingolani[1], Hugo Colombo[1], Pedro Hecht[2], Jorge Kamlofsky[1], Daniel Veiga[1].

[1]CAETI - Universidad Abierta Interamericana
Av. Montes de Oca 725, Buenos Aires, Argentina
abdel.masih@hotmail.com,{enrique.cingolani, hugo.colombo, jorge.kamlofsky, daniel.veiga}@uai.edu.ar

[2]Universidad de Buenos Aires, Facultades de Ciencias Económicas, de Ciencias Exactas y Naturales y de Ingeniería. Maestría en Seguridad Informática, Buenos Aires, Argentina.
phecht@dc.uba.ar

Resumen

Las computadoras cuánticas fueron creadas por primera vez en el año 1998 y actualmente empresas como Google, IBM y la NASA las están utilizando de modo experimental.

Su importancia radica en la velocidad con que procesan los datos, y lo hacen en un tiempo considerablemente menor que el llevado a cabo por una computadora convencional.

Por ejemplo, a un ordenador clásico le tomaría cientos de años poder encontrar los factores primos de un número natural de 600 dígitos, mientras que a una computadora cuántica sólo le tomaría unos minutos.

En este proyecto se estudian los fundamentos matemáticos de la Computación Cuántica.

Asimismo, se desarrollan algoritmos para ser implementados en un ordenador cuántico.

Si bien este nuevo paradigma de programación está en sus inicios, en un futuro no muy lejano las computadoras cuánticas reemplazarán a las convencionales en entornos de cálculos intensivos, y se tornará cada vez más necesario diseñar algoritmos acordes a esta nueva tecnología.

Palabras clave:
Computación Cuántica, Algoritmos Cuánticos, Quantum Computing, Compuertas Cuánticas, Qubits, Quantum Gates.

Contexto

Los proyectos desarrollados en el CAETI (Centro de Altos estudios en Tecnología Informática, dependiente de la Facultad de Tecnología Informática de la UAI) se clasifican en cinco líneas de investigación. Este proyecto se enmarca dentro la línea de investigación de "*Ingeniería de Software*". Se pretende brindar los fundamentos matemáticos y desarrollar algoritmos para ser utilizados en una computadora cuántica.

Introducción

Una computadora cuántica es un dispositivo que utiliza el modelo de los estados de ciertos elementos del átomo para realizar sus procesos.

Los ordenadores convencionales resumen toda la información que

procesan a lenguaje binario, es decir, sólo utilizan dos estados para almacenar y operar con datos: **0** ó **1**.

La unidad mínima de información utilizada por estos ordenadores es el ***bit***.

Pero en una computadora cuántica la unidad mínima de información es el ***bit cuántico*** o ***qubit***, el cual representa el estado físico de ciertos sistemas cuánticos, por ejemplo partículas subatómicas (que pueden ser electrones, protones, neutrones, fotones, etc.).

Dichos sistemas tienen una curiosa cualidad que hace fascinante a la Computación Cuántica: la ***superposición de estados***.

La ***superposición de estados*** en un sistema cuántico consiste en que éste no sólo puede adoptar el estado **0** ó **1**, sino que puede estar en ambos estados al mismo tiempo.

Gracias a este fenómeno surge el ***paralelismo cuántico***, que es la propiedad que tienen los algoritmos cuánticos de efectuar varias operaciones a la vez.

Así, los ordenadores cuánticos son capaces de probar, al mismo tiempo, todas las posibilidades que existen para la solución concreta de un problema, en lugar de probarlas una tras otra, como lo realizan actualmente las computadoras convencionales.

En definitiva, las computadoras cuánticas procesan los datos en forma paralela, a diferencia de las computadoras actuales, que lo hacen en forma secuencial [1].

Este cambio en el paradigma de la Computación supone un enorme salto hacia adelante, ya que permitirán realizar cálculos complejos que actualmente resultan inalcanzables en la computación clásica.

De este modo, podrán resolverse problemas clásicamente considerados "intratables" por su alto nivel de complejidad.

Gracias a la propiedad del paralelismo, las computadoras cuánticas pueden valuar, en forma simultánea, una función $f(x)$ para múltiples valores de la variable independiente x.

El Algoritmo de Deutsch- Josza [3], uno de los primeros algoritmos cuánticos, es un claro ejemplo que permite apreciar el funcionamiento y la potencia de los programas cuánticos mediante el uso del paralelismo cuántico.

Propuesto por David Deutsch y Richard Josza en 1992, su finalidad es determinar si una función booleana $f(x_1, x_2, \ldots, x_n)$ de n variables es constante (vale 0 o bien 1 en todas las entradas) o está balanceada (es decir, si toma el valor 1 para la mitad de las entradas y 0 para la otra mitad) [2].

Otro algoritmo cuántico muy importante en Ciencias de la Computación es el presentado por Lov Grover en 1997 [4] quien mostró cómo mediante una computadora cuántica se puede hallar un elemento dentro de una lista desordenada en $O(\sqrt{n})$ pasos, mientras que una computadora clásica lo logra en $O(n)$ iteraciones [5]. Es decir, el algoritmo de Grover permite hallar rápidamente resultados dentro de gigantescas bases de datos mediante una consulta simple. Esto es actualmente de gran importancia con el crecimiento de los sistemas de Big Data, alimentados por las nuevas tecnologías IOT (Internet de las Cosas) [6].

En los últimos años se han desarrollado otros algoritmos cuánticos con aplicaciones en múltiples áreas: Física, Química y Genética [7 – 10].

Sin embargo, el área donde la Computación Cuántica causó una revolución paradigmática es en la Criptografía.

Los algoritmos criptográficos asimétricos más usados del mundo [11 - 14] se basan en la dificultad para la

resolución de ciertos problemas numéricos conocidos como: IFP (Integer Factorization Problem) y DLP (Discrete Logarithm Problem). El paralelismo cuántico resulta ser una herramienta adecuada para afrontar estos problemas.

En 1997 Peter Shor presentó un algoritmo que reduce drásticamente la complejidad computacional del problema IFP mediante una computadora cuántica [15]. Los trabajos de Kitaev [16] y Proos-Zalka [17] presentaron ataques eficaces a los problemas DLP y DLP para curvas elípticas también mediante una computadora cuántica.

Hoy la existencia de la computadora cuántica es un hecho: la empresa D-Wave Systems ya vendió computadoras cuánticas a Lockheed Martin, al laboratorio Los Alamos, a Google y a la NASA, entre otros [19]. Además, IBM por su lado, ofrece servicios en la nube con su computadora cuántica [20].

En caso que los algoritmos [15 – 17] logren implementarse en computadoras cuánticas, hecho que la NIST (National Institute of Standards ando Technology) ve posible para un futuro cercano [21], arrasaría con la casi totalidad de los algoritmos empleados en la criptografía actual [18].

Líneas de Investigación, Desarrollo e Innovación

Se trabaja en dos ramas: Matemática y Algoritmos Cuánticos.

La Matemática es utilizada para modelar el comportamiento de los qubits, de las compuertas y de los circuitos cuánticos.

En la rama de los Algoritmos Cuánticos se pretende generar algoritmos para ser implementados en un ordenador cuántico. Estos son probados empíricamente mediante los servicios de computación cuántica ofrecidos por IBM en la nube [9].

Resultados y Objetivos

Como objetivo general, se pretende adquirir conocimientos teóricos y prácticos básicos acerca de la programación en computadoras cuánticas.

Los objetivos específicos son los siguientes:

■ Conocer acerca de las ventajas de la programación en computadoras cuánticas.
■ Investigar, analizar y diseñar circuitos lógicos basados en compuertas cuánticas.
■ Investigar, analizar y diseñar algoritmos diversos para computadores cuánticos.

Por otro lado, en el marco del proyecto se lograron los siguientes resultados a partir del año 2017:

Tesinas: En 2017 Pablo Oviedo presentó en la UAI su tesis para obtener el título de Licenciado en Matemática. La misma se titula: *Fundamentos matemáticos de computación cuántica en el algoritmo de Shor, para la factorización prima de números enteros*.

Dictado de Cursos: *Taller de Computación Cuántica*, en la XI Jornada de Matemática llevada a cabo en la UAI, el 9 de Mayo de 2018.

Premios obtenidos: Primer Premio en el Certamen de Trabajos Estudiantiles del CIITI 2017, que tuvo lugar en Buenos Aires, el 19 de septiembre de 2017. Título del Trabajo: *Algoritmo para la Factorización Prima de Números Enteros en una Computadora Cuántica*.

Diseño de algoritmos: Debido a que el proyecto está en sus inicios, se lograron probar solamente algunos algoritmos cuánticos básicos, con algunas modificaciones, en la computadora cuántica virtual de la empresa IBM. La misma constituye una herramienta de

suma utilidad, ya que es el único medio vigente que permite chequear los resultados de un algoritmo.

Formación de Recursos Humanos

El proyecto está dirigido por la Dra. Samira Abdel Masih y codirigido por el Lic. Jorge Kamlofsky. Integran el proyecto el Lic. Enrique Cingolani, el Ing. Hugo Colombo, el Dr. Pedro Hecht y el Lic. Daniel Veiga.

En el transcurso del año 2019 los estudiantes Yésica Valente y Ariel Savarese presentarán sus Tesis de la carrera de Licenciatura en Matemática, desarrollando temas referentes a algoritmos cuánticos.

Referencias

[1] Hecht, Juan Pedro. *Fundamentos de Computación Cuántica*. Editorial Académica Española. ISBN 978-3-8484-7529-2 (2005).

[2] Oviedo, Pablo. *Fundamentos Matemáticos de Computación Cuántica en el Algoritmo de Shor, para la factorización prima de números enteros* (Tesis de Grado). Universidad Abierta Interamericana, Buenos Aires, (2017).

[3] D. Deutch and R. Jozsa. *Rapid solution of problems by quantum computation.* Proc. Roy Soc. London Ser. pp 553-558. (1992)

[4] Grover, L. K. *Quantum computers can search arbitrarily large databases by a single query*. Phys. Rev. Lett. 79, pp 4709–4712 (1997).

[5] Spector, Lee. *Quantum Computing Applications of Genetic Programming.* Advances in Genetic Programming (1999).

[6] Gustafson S, and Sheth A. *Web of Things.* Computing Now 7.3 (2014).

[7] Holt, Katherine. *Diamond at the nanoscale: applications of diamond nanoparticles from cellular biomarkers to quantum computing*. Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences 365.1861 (2007): 2845-2861.

[8] Leuenberger, Michael N., and Daniel Loss. *Quantum computing in molecular magnets.* Nature 410.6830 (2001): 789.

[9] Cory, David G., Amr F. Fahmy, and Timothy F. Havel. *Ensemble quantum computing by NMR spectroscopy.* Proceedings of the National Academy of Sciences 94.5 (1997): 1634-1639.

[10] Steane, Andrew. *Quantum computing.* Reports on Progress in Physics 61.2 (1998): 117.

[11] Rivest, Ronald L., Adi Shamir, and Len Adleman: *A method for obtaining digital signatures and public-key cryptosystems.* Communications of the ACM 21.2, pp 120-126. (1978).

[12] Diffie W., Hellman M.E. *New directions in cryptography*. IEEE Transactions on information theory, 22, pp 644-654, (1976).

[13] El Gamal, Taher. *A Public Key Cryptosystem and a Signature Scheme Based on Discrete Logarithms. En Advances in cryptology.* Springer Berlin Heidelberg, pp. 10–18 (1984).

[14] Menezes, Alfred J., Paul C. Van Oorschot, and Scott A. Vanstone. *Handbook of applied cryptography*. CRC press (1996).

[15] Shor P. *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer.* SIAM J. Comput., 5, pp 1484-1509 (1997).

[16] Kitaev A. *Quantum measurements and the abelian stabilizer problem*. Preprintar Xiv/quant-ph., 9511026 (1995).

[17] Proos J., Zalka C. *Shor's discrete logarithm quantum algorithm for elliptic curves*. Quantum information and computation, 3, pp 317-344 (2003).

[18] Kamlofsky J, Hecht JP. *Post-Quantum Cryptography Using Hiper-complex Numbers.* XXIII Congreso Argentino de Ciencias de la Computación (2017).

[19] D-Wave-Systems Press Releases [en línea], (2018). Disponible en: <http://www.dwavesys.com/news/press-releases>. Fecha de consulta: 28/02/2018.

[20] IBM: IBM Makes Quantum Computing Available on IBM Cloud to Acceletare Innovation [En Linea], (2016). Disponible en: <https://www-03.ibm.com/press/us/en/pressrelease/49661.wss>. Fecha de consulta: 28/02/2018.

[21] National Institute of Standards and Technology, Information Technology Laboratory– Computer Security Division. *Post-Quantum Crypto Proyect* [En línea], (2018) Disponible en: <http: // csrc.nist.gov / groups / ST / post-quantum-crypto/>. Fecha de consulta: 28/02/2018.

Análisis de estrategias para clasificación de usuarios y post dentro de un hilo de discusión

Valeria Zoratto, Gabriela Aranda, Nadina Martinez Carod, Alejandra Cechich, 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: {vzoratto, gabriela.aranda, nadina.martinez}@fi.uncoma.edu.ar

RESUMEN

La Web actual se ha transformado en una plataforma que posibilita el encuentro de ideas y favorece la creación de debates en chat, blogs, foros de discusión, etc. En particular la comunidad informática suele aprovechar los medios disponibles en la Web de soporte grupal [1], tanto para solucionar problemas como para el aprendizaje de alguna tarea particular. Es por ello que este tipo de herramientas han tenido un gran auge en las últimas décadas, dentro de las cuales los foros de discusión se han convertido en los más utilizados para aprendizaje o como proveedor de soluciones de algún problema específico. Los foros de discusión generan contenido de manera continua lo que produce un gran volumen de información, que puede ser utilizado como fuente de conocimiento para un sistema de Information Retrieval (IR).

Las organizaciones actuales hacen cada vez más esfuerzos para reutilizar el conocimiento, definiendo estrategias para tener catalogadas y reutilizar soluciones ya probadas por lo que la disciplina de IR ha avanzado considerablemente.

El objetivo fundamental de nuestro proyecto es definir una herramienta que, a partir de información contenida en hilos de foros de discusión técnicos, pueda descargar dicha información de manera automática, la pueda clasificar de acuerdo a temas específicos, así como también poder establecer un ranking de soluciones posibles, teniendo en cuenta además a los usuarios involucrados en dichos foros.

CONTEXTO

La línea de investigación presentada se denomina “Reúso de Conocimientos en Foros de Discusión II” y forma parte del programa “Desarrollo de Software Basado en Reúso – Parte II”, con período de vigencia 2017-2020. El programa mencionado extiende el programa “Desarrollo de Software Basado en reúso” realizado durante el período 2013-2016.

1. INTRODUCCIÓN

La disciplina de Information Retrieval (IR) surge en la década de 1950 [2], ante la necesidad de reutilizar grandes volúmenes de información. 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.

Los foros son espacios web virtuales donde las personas pueden hacer preguntas, responder y participar en discusiones sin necesidad de estar en el mismo espacio geográfico ni en el mismo momento [3]. En los últimos tiempos han surgido foros

populares (específicos en el área de informática) como StackOverflow[1], Ubuntu Forum[2], etc. La información que se encuentra en dichos foros puede ser muy rica ya que puede ser utilizada por múltiples usuarios que tengan problemas similares.

La creación de un hilo comienza a partir de una serie de preguntas sobre un problema específico generada por un usuario de la comunidad del foro. Luego el usuario espera la respuesta del resto de los usuarios de la comunidad, estas respuestas pueden no ocurrir inmediatamente después de publicada la pregunta, ya que los participantes del foro no están físicamente presentes. De esta manera un usuario puede contestar a la pregunta estableciéndose así una comunicación entre las partes a fin de encontrar una solución.

Si bien las soluciones propuestas están focalizadas al usuario que abrió el debate, la totalidad de los mensajes queda disponible al público, y las soluciones pueden ser reutilizadas por participantes con problemas similares.

Muchos foros contienen diferentes tipos de usuarios, generalmente existen moderadores, administradores, usuarios registrados y usuarios anónimos. Algunos foros, asignan diferentes tipos de etiquetas a los usuarios registrados, dependiendo de la interacción que tenga este en la comunidad.

En general las técnicas actuales de IR se basan en medidas de similitud de palabra clave y no consideran algunas características importantes para analizar discusiones en hilos [4]. A menudo, los hilos que comparten palabras clave comunes discuten diferentes temas y, en tales casos, encontrar hilos relevantes se vuelve un desafío para los usuarios [5].

Este es un aspecto fundamental considerado en nuestro proyecto.

Existen varias propuestas de reúso de conocimiento disponible en foros de discusión: Por ejemplo, Elsas & Carbonell [6] fueron los primeros en revisar estrategias para la recuperación de hilos en una colección de prueba de pares <query, relevant document>. Por otro lado Seo et al. [7] describen cómo las estructuras de respuesta en los hilos de un foro pueden ser extraídos y utilizados para la recuperación a nivel de hilo y post. En otro trabajo, Bhatia & Mitra [3] emplean redes de inferencia para calcular evidencias de diferentes unidades estructurales. Este modelo utiliza múltiples atributos como el post inicial y las respuestas, longitud de los mensajes, medida de autoridad del post y los enlaces entre los posts en un marco de red de inferencia unificada basada en filas de clasificación en respuesta a las consultas de los usuarios.

Otros autores hacen foco en la jerarquía de información de los foros. Por ejemplo, Helic et al.[8], proponen clasificar los mensajes de foros de acuerdo a una jerarquía de temas pre-establecida. Luego, el enfoque de Nicoletti [9] clasifica los mensajes acordes a una jerarquía de temas obtenido de Wikipedia. Por otro lado, la investigación de Hecking et. at. [10] describe el análisis de la estructura social y semántica de los foros de discusión en cursos MOOC en términos de intercambio de información y roles de usuario.

Liu [11], en cambio, busca predecir si un autor de la pregunta estará satisfecho con las respuestas enviadas por los participantes de la comunidad.

En base a estos antecedentes, nuestro proyecto tiene como objetivo principal favorecer el reúso de la información contenida en conversaciones existentes en foros de discusión de la Web, con el valor agregado de un análisis de los roles de usuarios para poder detectar usuarios expertos y darle un mayor peso a las respuestas candidatas de una pregunta. 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

[1] http://stackoverflow.com
[2] https://ubuntuforums.org

análisis del lenguaje natural permite analizar el tipo de fragmento dentro de un hilo de discusión [12]. 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. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto de investigación se denomina "Reúso de Conocimientos en Foros de Discusión – Parte II" y está enmarcado dentro del Programa de Investigación "Desarrollo de Software Basado en Reúso – Parte II", con período de vigencia 2017-2020.

El programa mencionado extiende la investigación realizada durante el programa denominado "Desarrollo de Software Basado en Reúso", realizado en el período 2013-2016. Respecto a este proyecto en particular, el objetivo es extender los estudios realizados sobre reúso 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 frecuentes. Por otra parte, el programa está conformado por otros dos subproyectos que profundizan en las temáticas de Reúso Orientado al Dominio y Reúso Orientado a Servicios.

Dicho programa 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, junto con asesoría y colaboración de otras universidades. En particular, este proyecto es desarrollado en colaboración con 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 con un trabajo conjunto y colaborativo.

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 [13]. La validación de los atributos y sub-atributos de dicho modelo se realizó mediante encuestas [14]. 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 exhibidos en [15].

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 [16][17]. Para ello se utilizó la herramienta Lucene, con mecanismos personalizados para establecer stopwords (palabras que no aportan significancia) propias del dominio. En 2017, se aplicaron estas técnicas en combinación con la base de datos léxica WordNet [18], cuyos resultados preliminares fueron presentados en [19]. En 2018, se apuntó a mejorar el proceso de recuperación de hilos mediante la incorporación de sinónimos de las palabras utilizadas en los hilos de discusión teniendo en cuenta a la estructura semántica de la oración. Para ello se utilizó WordNet de

manera conjunta con la aplicación Stanford POS Tagger[3] [21]. 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 [27] 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 [22][23], así como técnicas y herramientas de PLN [24] 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). Bajo esta premisa, se han estudiado las propuestas [20] [25] [10] y se está trabajando en una tesina, a partir de una estrategia empírica basada en la observación de hilos de discusión obtenidos de la web. Por otro lado se están estudiando las propuestas de [11][28][29] y se está trabajando en otra tesis para detectar la satisfacción de una respuesta del usuario.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto avanza en la línea del proyecto comenzado en 2013, el cual tenía como objetivo definir un modelo de calidad a partir de información contenida en foros de discusión técnicos.

El proyecto actualmente se encuentra conformado por un grupo de docentes, asesores y alumnos desarrollándose en 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 Doctorado 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 Sistemas 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 el trabajo cooperativo 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. BIBLIOGRAFÍA

[1] C. A. Ellis, S. J. Gibbs, and G. L. Rein, "Groupware: Some Issues and Experiences", Communications of ACM 34, 1 (1991), pp. 38-58.

[2] Singhal, Modern information retrieval: A brief overview.IEEE Data Eng. Bull., 2001, vol. 24, no 4, p. 35-43

[3] S. Bhatia and P. Mitra, "Adopting Inference Networks for Online Thread Retrieval.", in AAAI vol. 10, (, 2010), pp. 1300--1305.

[4] D. Feng, E. Shaw, J. Kim, and E. Hovy, "Learning to detect conversation focus of threaded discussions", in Proceedings of the main conference on Human Language Technology Conference of the North American Chapter of the Association of … (, 2006), pp. 208--215.

[3] https://nlp.stanford.edu/software/tagger.shtml

[5] P. Biyani, S. Bhatia, C. Caragea, and P. Mitra, "Using Subjectivity Analysis to Improve Thread Retrieval in Online Forums", in European Conference on Information Retrieval (2015), pp. 495--500.

[6] J.L Elsas and J. G Carbonell, "It pays to be picky: an evaluation of thread retrieval in online forums", in Proceedings of the 32nd international ACM SIGIR (2009), pp. 714--715.

[7] J. Seo, W B. Croft, and D.A Smith, "Online community search using conversational structures", Information Retrieval 14, 6 (2011), pp. 547.

[8] D. Helic, N. Scerbakov (2003), "Reusing Discussion Forums as Learning Resources in WBT Systems".

[9] M. Nicoletti, S. Schiafino, and D. Godoy. Mining interests for user profiling in electronic conversations. Expert Syst. Appl., Feb. 2013.

[10] 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.

[11] Y. Liu, J. Bian, and E. Agichtein, "Predicting information seeker satisfaction in community question answering", 31st annual international ACM SIGIR (2008), pp. 483--490.

[12] 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.

[13] 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,

[14] N.Martínez Carod, G. Aranda. Análisis de la información presente en foros de discusión técnicos. In CACIC 2013, pp. 847- 856, 2013.

[15] 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, 2014.

[16] 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.

[17] V. Zoratto, G. Aranda, S. Roger, A. Cechich, Analyzing Discussion Forums ThreadsAbout Java Programming Language Usage, Electronic Journal of SADIO, 2016.

[18] 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.

[19] V. Zoratto, N. Martínez Carod, F. Otermín, G. Aranda: Análisis de estrategias para clasificar contenidos en foros de discusión, CACIC 2017, pp. 640-649

[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] G. Aranda, V. Zoratto, N. Martínez Carod, Sandra Roger, F. Otermín, A. Cechich. Clasificación de contenido de hilos de discusión mediante análisis sintáctico y morfológico. CICCSI 2018.

[22] I. Witten, E. Frank and M. Hall. Data Mining: Practical Machine Learning Tools and Techniques. Elsevier. 2011

[23] B. Liu. Web Data Mining. Exloring Hyperlinks, Contents, and Usage Data. Springer. 2008

[24] C. D. Manning and H. Schütze. Foundations of Statistical Natural Language Processing. MIT Press, Cambridge, MA, USA, 1999.

[25] 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

[26] G. A. Miller, R. Beckwith, C. D. Fellbaum, D. Gross, K. Miller. 1990. WordNet: An online lexical database. Int. J. Lexicograph.pp. 235–244.

[27] 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.

[28] Choi, E., Kitzie, V., & Shah, C. (2014). Investigating motivations and expectations of asking a question in social Q&A. *First Monday*, *19*(3).

[29] Agichtein, E., Liu, Y., & Bian, J. (2009). Modeling information-seeker satisfaction in community question answering. ACM Transactions on Knowledge Discovery from Data (TKDD), 3(2), 10.

Aplicación de la minería de procesos para determinar recursos computacionales en la ejecución de procesos de negocio

Diego Cocconi*, Marisa Pérez*, Juan Pablo Ferreyra*, Claudia Verino*

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

Resumen

La minería de procesos es una disciplina emergente que se puede usar para buscar y analizar la información que los sistemas de información registran sobre los procesos de negocio, para poder entenderlos, monitorearlos, analizarlos y mejorarlos. En este trabajo se propone buscar una mirada alternativa para la minería de procesos, utilizándola para estimar recursos computacionales necesarios para la implantación de la gestión de procesos de negocio (BPM, del inglés *Business Process Management*) en la organización. A tal fin, se propone emplear, extender y mejorar un simulador de procesos de negocio desarrollado anteriormente para otro proyecto.

Palabras clave: gestión de procesos de negocio, mejora continua, minería de procesos, simulación de procesos de negocio, sistemas de información.

Contexto

El presente trabajo se encuadra dentro del proyecto de investigación I+D UTN 4955 "*Implementación de técnicas de minería de procesos como herramienta de soporte y optimización de los procesos de negocio en la Secretaría de Extensión y Cultura de la U.T.N. 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 (UTN).

En el marco de dicho proyecto se propone como objetivo general mejorar los procesos en la Secretaría de Extensión Universitaria (SEU) de la UTN, Facultad Regional San Francisco, por medio de técnicas de minería de procesos para respaldar el modelo teórico validado en un proyecto anterior [1-2].

Durante el desarrollo de las actividades del proyecto actual se realizan además tareas de investigación que se desprenden del objetivo principal, como la que se describe en este trabajo.

1. Introducción

En la última década, la *gestión de procesos de negocio* (BPM, del inglés *Business Process Management*) se ha convertido en una disciplina madura que permite a las organizaciones administrar y mejorar sus procesos de negocio al aplicar un ciclo de mejora continua que se conoce como *ciclo de vida de BPM* (Figura 1), en el que intervienen las fases de *identificación de procesos*, *descubrimiento de procesos*, *análisis de procesos*, *diseño de procesos*, *configuración e implementación*, *ejecución y monitoreo*, y *evaluación* [3-5].

Las fases de identificación de procesos y descubrimiento de procesos producen como

resultado una *arquitectura* o *mapa de procesos*, que por lo general se compone de un conjunto de (macro) procesos de negocio que constituyen una vista global.

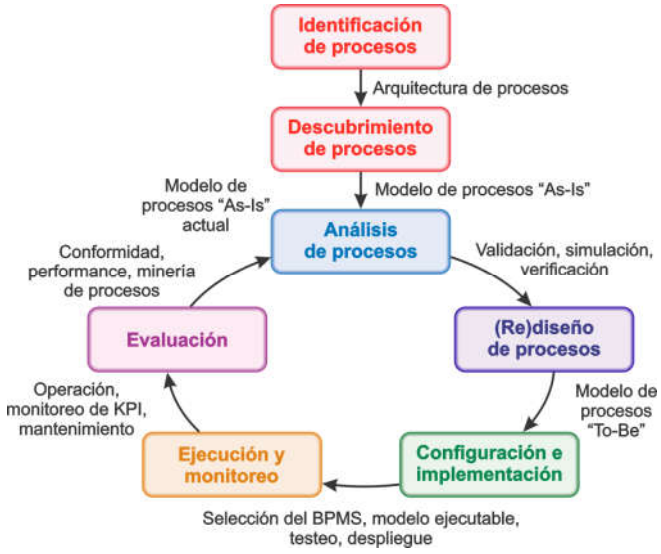


Fig. 1: *Ciclo de vida de BPM (adaptación de [3] y [4]).*

Durante el análisis se identifican cuestiones propias de los modelos documentados, incluyendo en lo posible *indicadores de performance* (del inglés *Key Performance Indicators*, KPI) [6]. La etapa de diseño identifica posibles cambios en los procesos que podrían mejorar los aspectos identificados durante el análisis. Durante la fase de configuración e implementación se configuran y especifican aspectos necesarios para que los modelos de procesos puedan ser interpretados y orquestados por un *sistema de gestión de procesos de negocio* (del inglés *Business Process Management System*, BPMS) [7], convirtiéndose en modelos ejecutables.

Luego, en la fase de ejecución y monitoreo, el BPMS crea instancias de los modelos de procesos para que puedan ser ejecutadas, haciendo uso de un *Sistema de Información Orientado a Procesos* (SIOP), del inglés *Process-Aware Information System*, PAIS) o de un *sistema de gestión de Workflows* (del inglés *Workflow Management System*, WfMS). Durante esta fase también se realiza el monitoreo de los KPI definidos, a fin de evaluar el desempeño. Finalmente, en la fase de evaluación se analizan los registros de ejecución de las instancias de procesos para aplicar ajustes necesarios a futuro [3], cuando se realice una nueva iteración del ciclo de vida.

Como sugiere el ciclo de vida, la mayoría de los enfoques de la gestión por procesos de negocio utilizan modelos que parten de una interpretación de la realidad para su posterior ejecución. Al no contemplar en esa interpretación los datos de hechos fácticos relacionados a los procesos puede existir una brecha entre el modelo teórico y la realidad [8]. En el mismo sentido, la generación de mejores modelos no necesariamente produce automáticamente mejores procesos [9]. Por lo tanto, los esfuerzos se deben alinear con el objetivo original de BPM de mejorar los procesos de negocio, en lugar de mejorar los modelos de proceso.

La minería de procesos es una disciplina emergente que se puede aplicar en la fase de evaluación para buscar y analizar la información que los sistemas de información registran sobre los procesos de negocio, para poder entender, monitorear, analizar y mejorar dichos procesos, mediante métodos y técnicas para proveer información basada en hechos y dar soporte a las mejoras de procesos [10].

El punto de partida para la minería de procesos es un *registro de eventos* (o *log*, en inglés). Los registros de eventos almacenan información de una actividad que está relacionada con un *caso* particular (una instancia de proceso) [11].

Existen tres tipos de minería de procesos (Figura 2): (1) *descubrimiento*, el cual toma un registro de eventos y produce un modelo sin utilizar ninguna otra información a priori; (2) *conformidad*, donde se compara un modelo de procesos existente con el registro de eventos del mismo proceso y esta comparación muestra dónde el proceso real se desvía del modelado; (3) *mejora*, que toma un registro de eventos y un modelo de procesos y extiende o mejora el modelo usando los eventos observados (es decir, cambia o extiende el modelo).

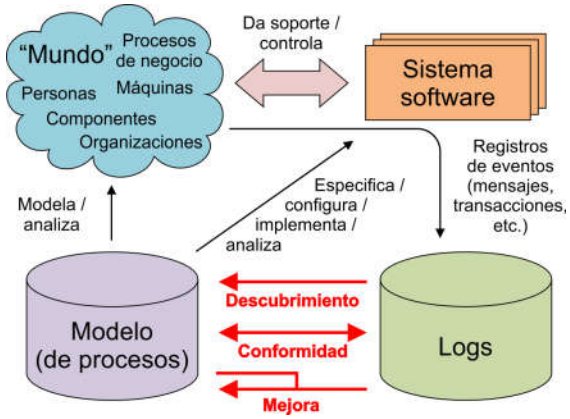


Fig. 2: *Los distintos tipos de minería de procesos [12].*

En este trabajo se propone buscar una mirada alternativa para la minería de procesos, asociada a una etapa preliminar de la implantación de BPM en la organización, pudiendo ser un complemento durante la fase de análisis.

A modo de caso de estudio, se considerarán como entrada los resultados de un proyecto I+D homologado anterior, el cual describe el trabajo de implantación de BPM en una secretaría de la UTN Facultad Regional San Francisco: la Secretaría de Extensión y Cultura [1-2].

2. Líneas de Investigación y Desarrollo

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.
- Minería de procesos.

Particularmente, en esta etapa de investigación y desarrollo se hace foco en la fase de evaluación de procesos (minería de procesos), considerándola un punto inicial para la etapa de análisis.

3. Resultados Obtenidos y Esperados

El principal objetivo del presente proyecto es utilizar minería de procesos para estimar recursos computacionales necesarios para la implantación de BPM en la organización.

Como primer paso se pretende realizar un análisis preliminar del contexto organizacional, determinando un bosquejo de la arquitectura de procesos y los procesos prioritarios (a tal fin se pueden emplear resultados de proyectos previos, como los descriptos en [1-2]).

A continuación, es necesario generar las instancias de proceso a partir de los modelos. No es necesario generar modelos ejecutables, se puede utilizar un simulador de procesos de negocio, como BIMP[1]. BIMP genera un archivo XML que registra los eventos producto de la simulación en forma de trazas, especificando los tiempos de ocurrencia.

El aspecto central constituye ahora en utilizar la información generada para determinar los recursos necesarios para la implementación de un SIOP o BPMS de acuerdo a la expectativa de ejecución de los procesos de negocio en la organización. Para ello se pretende utilizar un simulador de ejecución de procesos de negocio en la nube, el cual se encuentra desarrollando actualmente uno de los integrantes del proyecto en otra línea de investigación [13]. De este modo, se necesita extender dicho simulador para que considere la utilización de uno o pocos servidores, ya sea en la nube (teniendo en cuenta *balanceo de carga* y *escalamiento*) o no, físicos o virtuales (equivalentes a máquinas virtuales). La idea es ejecutar diferentes escenarios de configuración del(los) servidor(es) para verificar cómo se ven afectados los tiempos de ejecución de los procesos de negocio.

El trabajo de extensión incluye la mejora del proceso de importación del simulador, indagando algoritmos de minería de procesos que procesen la información proveniente del archivo XML de la simulación de procesos y

[1] http://bimp.cs.ut.ee/

realicen el mapeo hacia a la estructura de instancias necesarias para la simulación; particularmente, un aspecto de interés del caso de estudio considerado es la detección de paralelismos contando con escasas trazas de ejecución. La salida del algoritmo se ingresará a una estructura de colas que es la entrada del simulador, en la que se aplicará teoría de grafos para representar las dependencias de tareas que tienen como antecesoras varias tareas en paralelo y no pueden iniciar hasta que todas las instancias paralelas se hayan ejecutado (esto también representa una mejora del simulador).

Como resultado final, se espera que la simulación de la ejecución de los procesos de negocio bajo diferentes escenarios (que pueden variar en carga de trabajo, número de servidores, despliegue en la nube, etc.) permita determinar cuál configuración sería la más apropiada en términos de tiempos de ejecución de las instancias de procesos (y tareas) para el despliegue del SIOP.

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 Información. Entre los docentes, tres de ellos se encuentran en la etapa de desarrollo de sus tesis de maestría y otro está desarrollando actualmente su tesis de doctorado (mención Ingeniería en Sistemas de Información), todos ellos con temas altamente vinculados al área de estudio del proyecto.

Como iniciativa del grupo, se prevé la capacitación y formación de recursos humanos (cursos de posgrado, intercambio de ideas y conocimientos con personal de otras Facultades), 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, incorporación de la experiencia y los conocimientos obtenidos en el presente proyecto a las cátedras del tronco integrador de la carrera Ingeniería en Sistemas de Información, y el involucramiento de alumnos de la carrera Ingeniería en Sistemas de Información en la realización de actividades de este proyecto, incentivándolos a acercarse a propuestas de becas.

5. Referencias

[1] Perez, M., Ferreyra, J. P., Verino, C., Cocconi, D. (2017, Agosto). "Definición de una arquitectura de procesos utilizando la metodología BPTrends para la aplicación del ciclo de vida BPM". *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*.

[2] Pérez, M., Ferreyra, J. P., Verino, C., Cocconi, D. (2018). "Implementación de una arquitectura de procesos como resultado de la aplicación del ciclo de vida BPM durante sus fases de configuración y ejecución". *XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste)*.

[3] Weske, M. (2007). "*Business Process Management. Concepts, Languages, Architectures*". Springer-Verlag Berlin Heidelberg.

[4] Dumas, M., La Rosa, M., Mendling, J., Reijers, H.A. (2013). "*Fundamentals of Business Process Management*". Springer-Verlag Berlin Heidelberg.

[5] Cocconi, D., Ferreyra, J. P., Verino, C., Pérez, M. (2018, Noviembre). "Optimización organizacional basada en la aplicación del ciclo de vida BPM completo para la mejora continua de los procesos de negocio". *6to Congreso Nacional de Ingeniería en Informática / Sistemas de Información, Universidad CAECE*. Mar del Plata, Buenos Aires, Argentina.

[6] Harmon, P. (2014). "*Business process change*", 3era. ed. Morgan Kaufmann.

[7] 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".

5to. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI), 2017. Santa Fe, Argentina.

[8] Van Der Aalst, W. M. (2011, Abril). "Process mining: discovering and improving Spaghetti and Lasagna processes". *2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM) (pp. 1-7)*. IEEE.

[9] van der Aalst, W.M.P., La Rosa, M., Santoro F.M. (2016). "Business Process Management - Don't forget to improve the process!". *Business & Information Systems Engineering 58(1): 1-6, 2016*.

[10] Lazarte, I., Acosta Parra, C.A., Vilallonga, D.G. (2016). "Mejora y gestión de procesos de negocio inter-organizacionales aplicando técnicas de minería de procesos". *XVIII Workshop de Investigadores de Ciencias de la Computación., 2017., Bs.As., Argentina*.

[11] Van Der Aalst, W. M., Dustdar, S. (2012). "Process mining put into context". *IEEE Internet Computing, 16(1), 82-86*.

[12] van der Aalst, W. (2016). "*Process Mining. Data Science in Action*", 2nd ed. Springer-Verlag Berlin Heidelberg.

[13] Cocconi, D., Roa, J., Villarreal, P. (2019). "eBPSim: A Simulation Tool for Testing Elasticity Strategies in Cloud-based Business Process Solutions". *CIbSE 2019, La Habana, Cuba* (Aceptado, en prensa).

APLICACIÓN DE METODOLOGÍAS EMERGENTES PARA CONTRIBUIR CON LA CALIDAD DEL SOFTWARE

Irrazabal, Emanuel Dapozo; Gladys N.; Greiner, Cristina;
Company, Ana M.; Mascheroni, Agustín,

Facultad de Ciencias Exactas y Naturales y Agrimensura.
Universidad Nacional del Nordeste
{eirrazabal, gndapozo, cgreiner, eirrazabal, mascheroni }@exa.unne.edu.ar,
anamacom@gmail.com

RESUMEN

Este es el segundo año del proyecto F018-2017; una continuación de los proyectos F07-2009 y F10-2013, ambos enfocados en modelos, métodos y herramientas para la calidad del software. Este tercer proyecto aborda los temas emergentes en el área de la calidad de software, en particular, aspectos vinculados con la gestión de los procesos administrativos mediante herramientas software y en la entrega continua de productos software. El equipo de trabajo busca promover y/o generar métodos y herramientas que contribuyan a mejorar la calidad del proceso y del producto software.

En particular, se está trabajando en el desarrollo de procedimientos para la gestión de las actividades de prueba continua dentro de la disciplina de entrega continua. Esto ha incluido el desarrollo de revisiones sistemáticas de la literatura y el análisis comparado de las herramientas actuales de entrega continua y la evaluación de sus características. Finalmente también se atiende la problemática de la eficiencia de los sistemas en organismos públicos, mediante el enfoque de procesos de negocios y el desarrollo rápido de aplicaciones.

Palabras clave: calidad de software, gestión de procesos de negocio, 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

Las características de un producto software necesitan ser cada vez más complejas para satisfacer a los usuarios [1]. Esto genera avances constantes que se traducen en una gran cantidad de nuevas técnicas y tecnologías relacionadas con el desarrollo de software [2]. Pero el desafío sigue siendo llevar adelante un proceso de desarrollo que termine en un plazo acordado con unos atributos de calidad esperados, como, por ejemplo la mantenibilidad o la usabilidad.

En este sentido, la calidad del software está formada por diferentes factores y apoyados por herramientas que automatizan su medición. Por lo tanto, las nuevas tecnologías y técnicas emergentes muchas veces no tienen el conjunto de herramientas necesarias para la medición y mejora de la calidad del producto software resultante [3]. Por ello la necesidad de investigar en el uso y adecuación de las tecnologías emergentes incorporando aspectos de calidad del software.

Autores como Pressman [4], McCall [5], y estándares tales como la ISO 25010 [6] 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. Existen diferentes estudios que relacionan estos atributos con herramientas y métricas [7].

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 [8]. Pero en las tecnologías emergentes esto no está del todo desarrollado. Así por ejemplo, no se encuentran herramientas o técnicas de pruebas software para productos software nativos móviles [3].

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 [10] [11].

Nuevamente, 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 [11]. 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 [12].

Desde el enfoque de los procesos de construcción también existen desafíos, no solamente desde el sector privado sin también en el sector público. La gestión de procesos de negocio se basa en la idea de que cada producto es el resultado de un conjunto de actividades que se realizan a fin de obtener dicho producto. Por este motivo, la correcta y eficiente gestión de los procesos de negocio es un aspecto importante para la productividad de toda organización, ya que permite identificar las tareas, el orden de ejecución de las mismas y las personas responsables de realizarlas [13].

Por lo tanto, las líneas de trabajo del proyecto tienen que ver con características de calidad de calidad del producto software y del proceso de desarrollo.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

A continuación, se describen las siguientes líneas de investigación y desarrollo:

Entrega continua

Se está realizando el desarrollo de un conjunto de procedimientos para la gestión de las pruebas continuas teniendo en cuenta los principales desafíos de esta disciplina y los enfoques más actuales para agilizar el proceso de pruebas. En esta línea estamos especialmente orientados a las pruebas de tipo API Rest, las pruebas de visualización en navegadores y las pruebas de tipo flaky. Como primera medida se han realizado revisiones sistemáticas de la literatura y como segundo paso el desarrollo de encuestas a grandes empresas mundiales.

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. Y, finalmente, se propone implementar el modelo en un ambiente real de desarrollo, para analizar los resultados obtenidos.

Gestión de procesos de negocios 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.

En este sentido, la modelización rápida de prototipos para la automatización de las áreas administrativas es fundamental. Así, por ejemplo, una de las herramientas más utilizas es Bonita [14], una plataforma que integra tecnologías para el desarrollo de procesos de negocio. Esta herramienta además permite realizar entregas de aplicaciones operativas con rapidez y dispone de herramientas que posibilitan la mejora continua.

Incluye herramientas de modelización que siguen la notación BPMN 2.0 y contiene un motor de ejecución de procesos, o, por ejemplo, funcionalidades de integración para conectar las aplicaciones con servicios de terceros. Esto favorecerá la usabilidad del proceso [15].

En esta línea de investigación se busca realizar un modelo de trabajo para mejorar la gestión del proceso administrativo de las instituciones públicas. Por un lado, el modelado mediante la notación BPM y por otro lado el desarrollo de prototipos rápidos mediante la herramienta antes descripta.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el marco de este proyecto y respecto de la línea de gestión de procesos se lograron los siguientes objetivos:

1. Relevamiento de información relevante sobre los conceptos involucrados, procesos de negocios y usabilidad
   En primer lugar, se realizó una revisión sistemática de la literatura (RSL) con el objetivo de sintetizar la literatura existente sobre procesos de desarrollo de software basados en procesos de negocio y su vinculación con el diseño y la usabilidad de las aplicaciones. Las fuentes seleccionadas: IEEE Xplore, Science Direct, en el área de Computer Science y ACM Digital Library. Se encontraron pocos trabajos que estén estrechamente relacionados al objetivo de esta RSL, esto puede ser debido a que los procesos de negocio vinculados con el diseño de las aplicaciones y a la usabilidad es un tema poco desarrollado hasta el momento.
   El conocimiento extraído de esta RSL será la base para la construcción de un marco de trabajo para el desarrollo de software basado en BPM para la mejora de la usabilidad de aplicaciones [16].

2. Identificación y modelado de los procesos de negocio relevantes del área caso de estudio.
   Se seleccionó la plataforma BonitaSoft, que permite la creación de aplicaciones de negocio, mediante la definición del diagrama de procesos de negocio siguiendo la notación BPMN 2.0.
   Se seleccionó como caso de estudio el proceso de gestión de viáticos de empleados de un organismo estatal provincial.
   Utilizando BonitaSoft, que dispone de tecnologías para la automatización de los procesos de negocio, se continuará

con la digitalización del modelo del proceso en esta plataforma.
Además, se espera elaborar un marco de trabajo basado en el modelado de los procesos de negocio identificados y en técnicas de usabilidad.

En la línea de Ingeniería de Software para la Entrega Continua:

1. En [17] se resumieron e identificaron los principales problemas y desafíos que afectan a la calidad del software en las disciplinas de entrega continua y pruebas continuas teniendo en cuenta lo obtenido en revisiones sistemáticas de la literatura publicadas en años anteriores.
2. En [18] también se resumieron los factores claves para el éxito al realizar pruebas tipo flaky en servicios REST. Este tipo de pruebas tiene el problema de presentar diferentes resultados en cada ejecución. Por ello es necesario tener en cuenta un conjunto de buenas prácticas y estrategias, con el fin de disminuir sus consecuencias adversas.
3. En [19] se realizó una revisión sistemática de la literatura obteniendo las soluciones ante los problemas en pruebas continuas. Esta disciplina al ser nueva y estar enmarcada en las técnicas de entrega y despliegue continuo necesita ser adecuadamente analizada y revisada.
4. En [20] también se realizó una revisión sistemática de la literatura orientada al cross-browser test. Estos tipos de prueba son actualmente muy costosas, debido principalmente a la gran cantidad de configuraciones y dispositivos. Esto trae como consecuencia que un mismo sitio web sea visto de diferente manera en distintos navegadores. Muchas veces modificando su funcionalidad. Y es uno de los principales desafíos para las pruebas continuas, debido a este costo en esfuerzo a la vez que debe ser automatizado para lograr resultados beneficiosos.

En general, para esta línea de trabajo se espera desarrollar un modelo de buenas prácticas para la gestión de pruebas continuas en equipos de trabajo que tengan implantados la entrega continua de software. Este modelo hará especial hincapié en los tipos de prueba estudiados y detallados en esta sección.

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.

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] Randell, Brian. "Fifty Years of Software Engineering-or-The View from Garmisch." arXiv preprint arXiv:1805.02742 (2018).

[3] Kong, Pingfan, et al. "Automated testing of android apps: A systematic literature review." IEEE Transactions on Reliability 99 (2018): 1-22.

[4] R. Pressman , Ingeniería de Software. Un Enfoque Práctico, Sexta ed., MCGraw Hill, 2005.

[5] J. A. McCall, P. K. Richards y G. F. Walters, Factors in Software Quality, 1997.

[6] International Organisation for Standardisation, ISO/IEC 25010, Systems and software engineering -- Systems and software Quality Requirements and Evaluation

(SQuaRE) -- System and software quality models;Geneva, Switzerland, 2011.

[7] Arvanitou, Elvira Maria, et al. "A mapping study on design-time quality attributes and metrics." Journal of Systems and Software127 (2017): 52-77.

[8] R. Malhotra y A. Jain, «Software Fault prediction for Object Oriented Systems: A Literature Review,» ACM SIGSOFT Software Engineering Notes, vol. 36, nº 5, pp. 1-6 , 2011.

[9] Y. C. Cavalcanti, P. A. da Mota Silveira Neto y . 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

[10] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in IEEE Software 03/2015. V. 32(2).

[11] J. Humble and D. Farley. “Continuous delivery: reliable software releases through build, test, and deployment automation”, 1st ed. Boston, US: Pearson Education, 2010.

[12] 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

[13] Patricia Bazan. “Implementación de Procesos de Negocio a través de Servicios aplicando Metamodelos, Software Distribuido y Aspectos Sociales”. Tesis presentada para obtener el grado de doctor en ciencias informáticas. 2015.

[14] https://es.bonitasoft.com/plataforma-bonita

[15] Bicevskis, Janis, and Zane Bicevska. "Business process models and information system usability." Procedia Computer Science77 (2015): 72-79.

[16] Moreno, Juan Carlos, Marcelo Martin Marciszack, and Juan Pablo Fernandez Taurant. "Usability through Business Model." 2016 11th Iberian Conference on Information Systems and Technologies (CISTI). IEEE, 2016.

[17] Mascheroni, Maximiliano; Irrazábal, Emanuel. Problemas que afectan a la Calidad de Software en Entrega Continua y Pruebas Continuas. CACIC 2018

[18] Mascheroni, M. A., & Irrazábal, E. (2018). Identifying key success factors in stopping flaky tests in automated REST service testing. Journal of Computer Science and Technology, 18(02), e16-e16.

[19] Mascheroni, M. A., & Irrazábal, E. (2018). Continuous Testing and solutions for testing problems in Continuous Delivery: A Systematic Literature Review. Computación y Sistemas, 22(3).

[20] J Sabaren, L. N., Mascheroni, M. A., Greiner, C. L., & Irrazábal, E. (2018). A Systematic Literature Review in Cross-browser Testing. Journal of Computer Science and Technology, 18(01), e03.

Aplicaciones híbridas multiplataforma para Turismo Religioso

Susana I. Herrera[1], María M. Clusella[1], Silvia Sánchez Zuaín[1], Sergio H. Rocabado Moreno[2], David Cheein[1], Fernando Leturia[1], Stefano Trejo[1]

[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, mercedesclusella, silvisanzu, josedavidcheein, fleturia, stefanots}@gmail.com, srocabado@di.unsa.edu.ar

Resumen

Este artículo presenta avances y resultados del proyecto de investigación aplicada, que consiste en un desarrollo de una aplicación móvil multiplataforma sensible al contexto para el Turismo Religioso en Santiago del Estero Argentina, se enmarca dentro de la Ingeniería de software y aplicado al sector del Turismo. Es un Sistema de Información Turística basado en un Modelo Integral que permitirá, por un lado, la gestión de la Información por parte de la Dirección Provincial de Turismo y de los Proveedores de Servicios Turísticos y además contará con una aplicación móvil multiplataforma que le brindará al turista información en línea del contexto, sea cual fuere el SO de su dispositivo móvil.

Palabras clave: aplicaciones móviles multiplataforma, sensibilidad al contexto, turismo religioso.

1. Contexto

La investigación surgió a partir del análisis, de la situación actual de visitantes y de interesados en el ámbito del turismo religioso, realizado en investigaciones previas por UCSE, y por ende de la necesidad de contar con una herramienta Tecnológica de la Información y Comunicación, que responda a las exigencias de este tipo de turismo.

El equipo de investigación del proyecto está conformado por tres docentes investigadores de la Facultad de Ciencias para la Innovación y el Desarrollo (FCID) de la UCSE, un docente investigador de la Facultad de Ciencias Exactas de la Universidad Nacional de Santiago del Estero (UNSE), dos alumnos de posgrado y uno de grado de carreras de Informática de la FCID. Además se cuenta con el asesoramiento del equipo de investigación en Computación Móvil de la Universidad Nacional de Santiago del Estero.

2. Introducción

En los últimos años las aplicaciones móviles han constituido un ecosistema propio y un potente motor de innovación. La consolidación de las aplicaciones móviles como interfaz dominante del acceso a contenido tiene consecuencias en el plano de la producción, la distribución y el consumo, fusionando los cuatro ámbitos funcionales característicos del contenido digital: reproducción, creación, gestión y comunicación [1].

También es sabido que la aparición de los teléfonos inteligentes ha revolucionado multitud de campos; entre ellos, el del turismo. Como así también el uso de herramienta, conocidas como aplicaciones

informáticas ejecutadas desde teléfonos celulares (App), es una realidad latente e ineludible que forma parte de las dinámicas y procesos cotidianos de la vida; además de estar presente en la gestión de la información de toda índole (personal, académica, laboral, de entretenimiento, etc.) [2].

Es por ello que los sistemas móviles son uno de los pilares fundamentales en el fomento del turismo en muchos lugares del mundo, en particular, los sistemas sensibles al contexto, en los cuales, la ubicación del usuario, es la principal variable. A partir de ésta, los sistemas brindan diversos servicios al usuario, teniendo en cuenta su contexto social y sus preferencias.

Gracias a los elementos de hardware incorporados hoy en día en los Smartphones, como lo son el GPS, la tarjeta de red, los sensores, etc; sumado al acceso casi permanente a Internet sea por medio de una red móvil o una red inalámbrica y a la información personal que comparte el usuario, las aplicaciones móviles pueden ofrecer servicios y productos específicos que satisfagan sus necesidades [3].

Cada vez es más la demanda del Turismo religioso en la Provincia de Santiago del Estero. Por lo tanto, sería muy útil para las personas, que realizan visitas turísticas de tipo religioso, contar con una aplicación móvil que en lugar de enfocarse en informar sobre la ubicación de negocios locales o alojamiento, ofrezca información útil sobre los atractivos y puntos de interés relacionados con la religión.

Esta investigación aplicada involucra el desarrollo de una aplicación móvil multiplataforma sensible al contexto. La aplicación brindará al turista, sea cual fuere el SO de su dispositivo, información del entorno en línea poniendo a disponibilidad de ellos información útil sobre destinos turísticos, hoteles, servicios, entre otros; 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, y así maximizar su oferta a los potenciales turistas logrando mayores ingresos y competitividad. Para solucionar esto, Herrera et al. [4], proponen un modelo integral y colaborativo de turismo, denominado "Marco Sistémico para el Desarrollo de Aplicaciones de m-Turismo", que gestiona dicho contenido dinámico.

Los resultados que se presentan en este trabajo, se considera que 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

La línea de investigación y desarrollo de este trabajo tiene como eje central, el uso de herramientas de desarrollo híbrido que permitan desarrollar eficientemente aplicaciones móviles multiplataforma sensibles al contexto.

La hipótesis propone una dependencia de la segunda sobre la primera variable. Esta relación será contrastada mediante un desarrollo experimental.

El mismo involucra la definición de indicadores para la variable desarrollo eficiente y el desarrollo de prototipos que permitan evaluar la eficiencia.

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.

Para poder dar cumplimiento a estos objetivos se plantearon los siguientes objetivos específicos:

- 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 [4] 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
- Analizar la eficiencia de los prototipos generados para los SO Android y iOS.

A continuación se presentan los resultados obtenidos hasta la fecha.

Se realizó un análisis comparativo de 3 herramientas de desarrollo híbrido de aplicaciones móviles: Apache Cordova, Ionic y PhoneGap. Se definieron aspectos a considerar, en la comparación, relacionados con la eficiencia de las aplicaciones que generan; entre ellos: tecnologías que usan, entornos de desarrollo, licenciamiento, costos, servicios ofrecidos, soporte, documentación, plataformas para las cuales generan las aplicaciones.

En función de este análisis, se decidió desarrollar el prototipo usando Apache Cordova [3].

En [3] se reportó el desarrollo de un pequeño prototipo usando Apache Cordova 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 Apache Cordova.

En una segunda etapa, la aplicación generada fue probada en dispositivos Android. La prueba se realizó con un dispositivo Motorola Moto X2 con la versión 5 del sistema operativo. Con la ubicación activada, en un lugar aleatorio, se procedió a ejecutar primero la aplicación desarrollada en Apache Cordova.

A modo comparativo, se verificaron las ubicaciones utilizando dos apps nativas llamadas “Mi ubicación GPS” y “Google Maps”.

Se realizaron diversas pruebas, tanto en lugares “in-door” como “out-door”. En todos los casos los resultados tuvieron pocos metros de diferencia (+/- 5 mts), según la representación en el mapa. Por lo tanto, se puede considerar que el funcionamiento del GPS en una aplicación híbrida es el mismo al de una aplicación nativa de Android.

Se realizó relevamiento de conectividad en zonas aisladas de turismo religioso (escasa conectividad y recursos energéticos limitados). Por otra parte, se realizó relevamiento de eficiencia de equipo amplificador de red de datos. Ello permitió elaborar un Modelo de conectividad para zonas aisladas.

En cuanto al relevamiento de conectividad, el mismo se llevó a cabo en el circuito de turismo religioso "Camino Real", abarcando capillas y santuarios de: San Pedro, Upianita, Manogasta, Tuama, Villa Silípica, Sumamao, Loreto y Atamisqui. Localidades de la Provincia Santiago del Estero.

En cada punto, se efectuaron mediciones utilizando las aplicaciones:

• Network Cell Info: provee la ubicación de la señal en el mapa y permite medir la potencia de la red para señal y señales vecinas.

• METEOR- test de velocidad: se utiliza para para medir la velocidad de subida y bajada de archivos, usando aplicaciones móviles típicas como Whatsapp, o YouTube.

En relación al relevamiento de la eficiencia y consumo de energía de la antena amplificadora de señal, se realizó en zonas rurales aisladas de la Provincia de Salta. A partir de los resultados obtenidos [5], se diseñó el modelo de conectividad propuesto para el uso en puntos de interés aislados (nodos de las rutas de turismo religioso).

Se realizó la revisión y la adaptación del modelo de turismo de Herrera et al. [4], el cual prevé que los servidores con la información deben estar centralizados y a cargo de un ente gubernamental responsable del Turismo en la provincia.

Se realizó el desarrollo de la aplicación móvil siguiendo la metodología AUP (Agile Unified Process). La metodología AUP permite prototipación rápida. Que a su vez permite refinamiento del prototipo.

A partir de la definición de requisitos preliminar se advirtió la necesidad de desarrollar además de la aplicación móvil una Progressive Web Apps o PWA [6], para la gestión de la información por parte de la Dirección Provincial de Turismo y de los proveedores de servicios turísticos, que son básicamente aplicaciones web que ofrecen una experiencia de usuario muy cercana a la de una aplicación móvil nativa.

Se avanzará, cuando se obtenga el prototipo final, con el análisis de la eficiencia de los prototipos generados para Android y iOS.

5. Formación de recursos humanos

Dos integrantes del equipo, están finalizando sus trabajos finales de la carrera Especialización de Ingeniería web de la UCSE. La directora realizó estudios de posdoctorado referido aplicaciones móviles con interfaces con RA en la Universidad de las Islas Baleares (España).

6. Referencias

[1] Aguado, J. M.; Martínez, I. J.; & Cañete-Sanz, L. (2015). Tendencias evolutivas del contenido digital en aplicaciones móviles. El profesional de la información; 24(6), 787-795.

[2] Bohórque Garzón, D.; Chaviano,O.(2017)Implementation of mobile applications for the management of research from bibliometric information Anales de Investigación; 13(2), 158-168.

[3] Herrera, S.; María M. Clusella, M.M;Sánchez Zuaín,S.; Rocabado, S.Cheein,D.; Leturia, F.; Trejo S (2018). Aplicación móvil multiplataforma sensible al contexto para el turismo religioso en Santiago del Estero. Congreso CACIC Facultad de Ciencias Exactas – Tandil.

[4] Herrera, S.; Najar Ruiz, P.; Contreras, N.; Fennema, C.; Lara, C. Marco Sistémico para el Desarrollo de Aplicaciones de m-Turismo. Libro. Artículo Completo. Congreso. Congresso Brasileiro de Sistemas: pensar a compreensão de sistemas. Universidade Federal do Tocantins, 2013. Brasil. Palmas, Tocantins. ISBN 978-85- 89102-43-8.

[5] Herrera, S; María M. Clusella, M.M;Sánchez Zuaín,S; Rocabado,

S;Cheein,D.; Leturia, F.; Trejo S (2018)Sistema de Información Turística basado en un Modelo Integral. Congreso CONAIISI-Universidad CAECE – Mar del Plata, Buenos Aires.

[6] Progressive Web Apps. Google Developers. Disponible en: https://developers.google.com/web/progressive-web-apps/

Aplicaciones Móviles 3D Inmersivas

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 tridimensionales (3D) sobre diversas plataformas. En particular, se pone el foco en el análisis de performance y consumo de energía de las aplicaciones 3D, como así también en el desarrollo de aplicaciones 3D inmersivas, a través del uso de Realidad Virtual (RV).

Palabras claves: Dispositivos Móviles - Aplicaciones 3D - Aplicaciones Multiplataforma - M-Learning – Performance – Consumo de energía – Realidad Virtual

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.

La RV es una simulación interactiva por computadora en la cual se sustituye el mundo real con información sensorial que recibe el usuario. La RV permite generar entornos inmersivos donde el usuario puede interactuar con representaciones virtuales de objetos, que de otro modo sería difícil o imposible de acceder. De esta forma, es posible desarrollar aplicaciones donde el usuario se vincule con elementos históricos, obras artísticas u otros, que tienen un acceso físico nulo o muy limitado [1].

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 [2] [3] [4].

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 [5].

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 [6] [7].

Líneas de Investigación y Desarrollo

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para dispositivos móviles.
- Aplicaciones Móviles 3D Multiplataforma
- Mobile Learning
- Frameworks para el desarrollo de Aplicaciones Móviles 3D
- Performance de aplicaciones móviles 3D
- Consumo de energía en aplicaciones móviles 3D
- Realidad Virtual 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 [8] [9] [10].
- 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 ejecución de aplicaciones generadas con diferentes frameworks 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 ampliado el prototipo móvil InfoUNLP3D [11], que sirve de guía para los estudiantes en sus primeras experiencias dentro de la Facultad de Informática. El prototipo ahora puede ser controlado mediante lentes de RV, permitiendo una inmersión completa en el escenario virtual. Figura 1.
- Se ha desarrollado mediante Unity Engine un prototipo de análisis de performance de ejecución de aplicaciones 3D [12]. Figura 2.

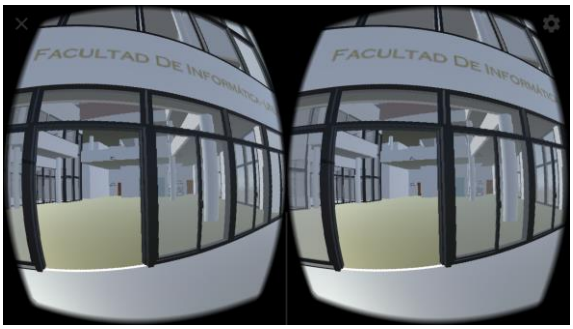


Figura 1. InfoUNLP3D controlado mediante las lentes de RV.

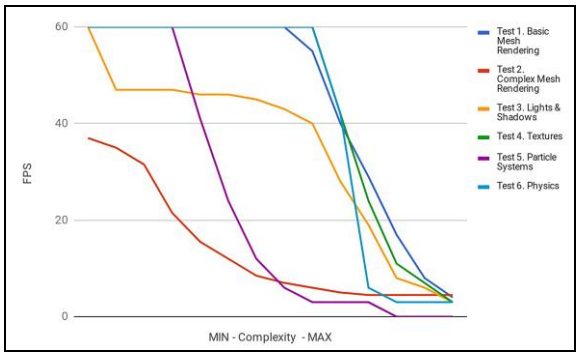


Figura 2. Prototipo de análisis de performance de ejecución. Resultados obtenidos.

- Se ha desarrollado mediante Unity Engine un prototipo de análisis de consumo de energía de las aplicaciones 3D [13]. Figura 3.

- Se está avanzando en el desarrollo de una aplicación móvil 3D inmersiva que permita recorrer el barco en el que Charles Darwin viajó a la Provincia de Buenos Aires para realizar diversas exploraciones. La aplicación permitirá movilizarse por la embarcación y visualizar sus descubrimientos mediante lentes de RV. Figura 4.

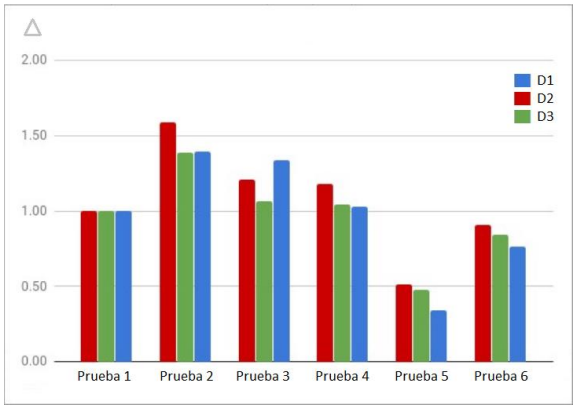


Figura 3. Prototipo de análisis de consumo de energía. Resultados obtenidos.

Figura 4. Modelo 3D de la embarcación que utilizó Darwin en su viaje a la Provincia de Bs.As.

- Se está avanzando en el desarrollo de un prototipo de análisis de performance de las aplicaciones 3D desarrolladas mediante el framework Unreal Engine con el fin de realizar un estudio comparativo con el prototipo desarrollado en Unity.
- Se está avanzando en el desarrollo de un prototipo de análisis de consumo de energía de las aplicaciones 3D desarrolladas mediante el framework Unreal Engine con el fin de realizar un estudio comparativo con el prototipo desarrollado en Unity.

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. Linowes J. “Unity Virtual Reality Projects”. 2015. ISBN-13: 978-1783988556.
2. 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.
3. 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.
4. 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.
5. Akekarat Pattrasitidecha. "Comparison and evaluation of 3D mobile game engines". Chalmers University of Technology. University of Gothenburg. 2014.
6. Riaz, M. "Energy consumption in hand-held mobile communication devices: A comparative study". International Conference on Computing, Mathematics and Engineering Technologies (iCoMET). 2018. ISBN: 978-1-5386-1370-2/18.
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. Unity 3D Homepage: https://unity3d.com/.

9. Unreal Engine Homepage: https://www.unrealengine.com/.
10. CryEngine. Homepage: https://www.cryengine.com/
11. 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.
12. 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.
13. Federico Cristina, Sebastián Dapoto, Pablo Thomas, Patricia Pesado. "Análisis de consumo de energía en aplicaciones 3D sobre dispositivos móviles". CACIC 2018. Tandil, Argentina.

Aplicaciones para Dispositivos Móviles. Estrategias y Enfoques de Desarrollo.

Pablo Thomas, Lisandro Delia, Leonardo Corbalan, Germán Cáseres, Juan Fernandez Sosa, Fernando Tesone, Verónica Aguirre, Verena Olsowy, 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)

526 e/ 10 y 11 La Plata Buenos Aires

{pthomas, ldelia, corbalan, gcaseres, jfernandez, ftesone, vaguirre, volsowy, 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 y Latinoamérica.

1. INTRODUCCIÓN

Los avances tecnológicos en la comunicación móvil, junto con los avances en el campo de la computación han dado lugar a la computación móvil, que puede definirse como un entorno de cómputo con movilidad física que brinda al usuario la capacidad 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 [1].

Las características específicas de la computación móvil plantean nuevos desafíos en el desarrollo de software para este tipo de dispositivos. La necesidad de tratar con diversos estándares, protocolos y tecnologías de red; las capacidades limitadas, no sólo en cuanto a procesamiento, sino también, en cuanto a consumo, tamaño físico y capacidad de almacenamiento 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, a veces para brindar movilidad a una parte de la funcionalidad de un sistema más grande, mientras que otras veces son el único punto de interacción del sistema.

Algunas iniciativas como *offline first* proponen nuevas maneras de abordar el desarrollo de aplicaciones móviles tomando como requerimiento no funcional la posibilidad de seguir brindando un servicio incluso cuando el dispositivo pierde conectividad con la red móvil.

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 específico de aplicaciones de manera nativa en cada una de las plataformas existentes, utilizando entornos de desarrollo, lenguajes de programación y tecnologías propias de cada plataforma.

El desarrollo nativo de aplicaciones móviles posee ciertas ventajas como la posibilidad de acceder sin limitaciones a todas las características del dispositivo (cámara, GPS, acelerómetro y agenda, entre otras), el alto rendimiento, interfaces gráficas consistentes con el resto de la plataforma, la posibilidad de trabajar sin acceso a Internet y de correr procesos en segundo plano. 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 reutilizar el código fuente entre plataformas diferentes, el esfuerzo se multiplica y se elevan los costos de desarrollo, actualización y distribución de nuevas versiones.

Como alternativa al desarrollo nativo en múltiples plataformas, se presentan diversos enfoques de desarrollo que procuran optimizar la relación costo/beneficio compartiendo la misma base de código fuente entre las versiones para las distintas plataformas [2]. 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 tecnologías web bien conocidas (HTML, CSS y JavaScript), no necesitan adecuarse a ningún entorno operativo móvil específico; 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, la experiencia de usuario dista de ser similar a la de las aplicaciones nativas, resultando menos atractiva para el usuario final.

El enfoque híbrido es una alternativa que posee también la ventaja de estar basado en tecnologías web estándar (HTML, Javascript y CSS) pero, a diferencia del anterior no funciona dentro de un navegador, sino en un contenedor web especial con mayor acceso a las características 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.

El enfoque interpretado se presenta como una alternativa en donde las aplicaciones 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 Javascript, Typescript y XML, entre otros.

La obtención de interfaces nativas constituye una de las principales ventajas de este tipo de aplicaciones, mientras que la definición de nuevas componentes suele tener un alto grado de complejidad debido a la necesidad de definir abstracciones compatibles con diferentes plataformas.

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.

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.

2. 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 en Android [3].
- Aplicaciones Nativas en Ios [3].
- Aplicaciones Web Móviles.
- Aplicaciones Móviles Híbridas (PhoneGap [4], Ionic [5]).
- Aplicaciones Móviles Interpretadas (Appcelerator Titanium [6], NativeScript [7]).
- Aplicaciones Móviles generadas por compilación cruzada (Xamarin [8], Corona [9]).
- PWA, Instant App y Offline First.
- Análisis y estudio comparativo de requerimientos no funcionales, tales como 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.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados obtenidos/esperados se pueden resumir en:

- Se ha realizado un estudio [10] sobre 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 elaborado un detallado estudio comparativo [11] que contempla una lista de 23 características de interés para la Ingeniería de Software y su impacto en 9 tecnologías para el desarrollo de aplicaciones móviles.
- Se ha presentado un minucioso estudio sobre la eficiencia energética y su relación con los enfoques de desarrollo [12]. Se estudió 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.
- Se ha estudiado [14][15] 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 utilizando varios frameworks de desarrollo multiplataforma.
- Se continúa con el desarrollo y evolución de "*Informática UNLP*", una aplicación móvil [13] 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. La app incluye una cartelera virtual donde los alumnos reciben información de interés sobre sus cursos. Además mediante realidad aumentada permite consultar información de la ocupación de las aulas en tiempo real.
- 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.
- Analizar las ventajas y desventajas de las *"Instant App"* de Android. Resulta de interés su estudio y posterior comparación con las aplicaciones nativas tradicionales.
- Promover el avance sostenido y continuo de la formación de los miembros involucrados en ella.

4. 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 este proyecto. Además participan en el dictado de asignaturas/cursos de grado y postgrado de la Facultad de Informática de la UNLP.

5. BIBLIOGRAFÍA

1. Hongxing Li, Guochu Shou, Yihong Hu, Zhigang Guo. Mobile Edge Computing: Progress and Challenges. 2016 4th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering (MobileCloud). Oxford UK.

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. http://phonegap.com/
5. https://ionicframework.com/
6. http://www.appcelerator.com/
7. https://www.nativescript.org/
8. http://xamarin.com/
9. https://coronalabs.com/
10. J. Fernández Sosa, P. Thomas, L. Delía, G. Cáseres, L. Corbalán, F. Tesone, A. Cuitiño, P. Pesado. ***"Mobile Application Development Approaches: A Comparative Analysis on the Use of Storage Space"***. XV Workshop de Ingeniería de Software - XXIV Congreso Argentino de Ciencias de la Computación CACIC 2018. Tandil, Argentina. 8 al 12 de octubre de 2018.
11. L. Delia, P. Thomas, L. Corbalan, J. Fernandez Sosa, A. Cuitiño, G. Cáseres, P. Pesado. ***"Development Approaches for Mobile Applications: Comparative Analysis of Features"*** SAI - Computing Conference 2018. London, United Kingdom. 10 al 12 de Julio de 2018.
12. L. Corbalan, J. Fernandez Sosa, A. Cuitiño, L. Delía, G. Cáseres, P. Thomas, P. Pesado. ***"Development Frameworks for Mobile Devices: A Comparative Study about Energy Consumption".*** 5th IEEE/ACM International Conference on Mobile Software Engineering and Systems. Gothenburg, Sweden. 27 y 28 de Mayo del 2018.
13. 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.
14. 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.
15. 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.
16. 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
17. 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
18. Digital Possibilities. Mobile Development Frameworks Overview http://digital-possibilities.com/mobile-development-frameworks-overview/
19. Markus Falk. Mobile Frameworks Comparison Chart, http://www.markus-falk.com/mobile-frameworks-comparison-chart/
20. Anup Kumar y Bin Xie, Handbook of Mobile Systems Applications and Services. Editorial CRS Press, ISBN 978-1-4398-0152-9, Año 2012.
21. 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).
22. 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)

23. Anthony Wasserman , Carnegie Mellon Silicon Valley, Software Engineering Issues for Mobile Application Development, 2º Workshop on Software Engineering for Mobile Application Development, MobiCASE ’11, Santa Monica, California, USA, October 2011.

24. 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.

25. Tim A. Majchrzak, Andreas Biørn-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

Aportaciones al proceso de implantación de sistemas informáticos

Marisa Panizzi[1,2], Mauricio Davila[1], Agustin Hodes[1], Pablo Vázquez[1], Felipe Ortiz[1], Rodolfo Bertone [2], Alejandro Hossian[3]

[1] Departamento de Ingeniería en Sistemas de Información. Universidad Tecnológica Nacional. Facultad Regional Buenos Aires.
Medrano 951 (C1179AAQ), C.A.B.A, Argentina.
2 Instituto de Investigación en Informática (III-LIDI). Facultad de Informática.
Calles 50 y 120 - La Plata - Bs. As. - Argentina

[3] Grupo de Investigación en Aplicaciones de Sistemas Inteligentes en Ingeniería. Universidad Tecnológica Nacional. Facultad Regional Neuquén.
Av. Pedro Rotter S/N Barrio Uno, (8318), Neuquén, Argentina

marisapanizzi@outlook.com, davilamr.80@gmail.com, hodesagustin@gmail.com, vazpablo@gmail.com, ortizfd@gmail.com, pbertone@lidi.info.unlp.edu.ar, alejandrohossian@yahoo.com.ar

Resumen

Esta línea de investigación tiene por finalidad realizar aportaciones paliativas al proceso de implantación de sistemas informáticos. La aplicación de métodos de investigación de Ingeniería de Software Basada en Evidencia y métodos de evaluación, permitió evidenciar que varias de las metodologías de desarrollo existentes no contemplan el proceso de implantación de un sistema informático o, si lo hacen, no lo desarrollan o gestionan de manera integral. En respuesta a estas debilidades se proponen diferentes aportaciones.

La primera consiste en un modelo de proceso de implantación de sistemas informáticos, para su aplicación en la industria del software en el ámbito nacional.

La segunda se corresponde con un conjunto de métricas específicas para el proceso en estudio, con el propósito de mejorar su nivel de calidad.

Y por último, a partir de lograr evidenciar la ausencia de riesgos para el proceso de implantación, se considerará una propuesta de riesgos y sus procedimientos de mitigación.

Se realizarán pruebas de concepto en los casos de estudio y casos de validación identificados que corroboren las propuestas de esta línea de investigación.

Palabras clave: procesos de software, implantación de sistemas informáticos, métricas, riesgos, modelo de proceso.

Contexto

La investigación que se reporta en este artículo es financiada parcialmente por el Proyecto UTN4347 de la Secretaria de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional.

La línea de investigación que se presenta en este proyecto, cuenta con el asesoramiento científico del grupo de investigación de Ingeniería de Software del Instituto de Investigación en Informática (III-LIDI) de la UNLP.

Introducción

La constante evolución de la ingeniería de software ha logrado un conjunto de metodologías de desarrollo o estándares para la construcción de sistemas informáticos como así también modelos de mejora de procesos de software.

En esta línea de investigación, se busca fortalecer el proceso de implantación de sistemas informáticos. En el contexto de esta línea de investigación, se considera al proceso de implantación, como el conjunto de actividades y tareas necesarias que permiten la transferencia del producto software finalizado a su ambiente de utilización por parte de la comunidad usuaria (Panizzi M. et al. 2017).

Con el propósito de identificar, evaluar e interpretar toda la investigación disponible relevante respecto al proceso de implantación de sistemas informáticos, se han utilizados métodos de investigación propios de la Ingeniería de Software. Entre ellos, el mapeo sistemático de la literatura (en Inglés Systematic Mapping Studies o SMS), es uno de los métodos de Ingeniería de Software Basada en Evidencia (EBSE), propuesto por Bárbara Ann Kitchenham. Este método permite identificar, evaluar e interpretar toda la investigación disponible para responder a unas preguntas de investigación. (Kitchenham B. et al., 2007). El SMS es un medio imparcial y repetible de identificar, evaluar e interpretar toda la investigación disponible relevante para una pregunta de investigación, área temática o fenómeno de interés (Kitchenham, B. et al., 2007), (Genero M. et al., 2014).

El otro método empleado, es el método DESMET (Kitchenham B., 1996), del cual se ha seleccionado el método de evaluación cualitativa basada en el análisis de características, con las adecuaciones necesarias a los fines de esta investigación.

El primer aporte en el que se está trabajando, es la construcción de un modelo de proceso que permita sistematizar el proceso de implantación de software. Este modelo podrá ser utilizado en conjunto, con la metodología de desarrollo de la organización productora de software. Este modelo de proceso determina fases, actividades, tareas, técnicas y prácticas, roles y las competencias de cada uno de los roles.

El segundo aporte al proceso de implantación de sistemas informáticos, es definir un conjunto de métricas.

El tercer aporte de esta línea para fortalecer el proceso en estudio, es la propuesta de un conjunto de riesgos a considerar como así también los procedimientos necesarios para su mitigación.

Para la construcción del conjunto de métricas y la propuesta de los riesgos, el grupo de investigación, considera adecuada una visión del proceso de implantación, en tres dimensiones. La primera dimensión denominada: ¨*Proceso*¨ dado que resultan de interés las fases o etapas, actividades y tareas que lo componen. La segunda dimensión denominada ¨*Producto*¨, la cual contempla las características como la complejidad del producto a instalar, los requisitos de instalación para el producto software, la integración con la infraestructura del cliente, entre otras. La última dimensión se denomina ¨*Persona*¨, debido a la existencia del peopleware y su impacto en el proceso en estudio (Panizzi M. et al., 2016). Esta visión tridimensional del proceso, se basa en la propuesta de clasificación de métricas considerada por Vázquez et al. (Vázquez P. et al., 2018).

Los autores consideramos que en los proyectos de software, generalmente se resta importancia al proceso de implantación o entrega del sistema informático por tratarse de uno de los eslabones finales de la cadena productiva del producto software (Panizzi M. et al, 2016).

El grupo de investigación, se plantea las siguientes preguntas de investigación (PI):

PI1: ¿Se puede lograr cubrir la vacancia de un Modelo de Proceso para la implantación de un sistema informático, que integre las actividades y sus tareas, las técnicas, las herramientas, los artefactos, los roles y sus competencias?

PI2: ¿Se puede mejorar la calidad del proceso de implantación mediante la determinación de métricas específicas para el mismo?

PI3: ¿Se puede fortalecer el proceso de implantación mediante la identificación de un conjunto de riesgos como así los procedimientos para su mitigación?.

El desafío de esta línea de investigación, consiste en lograr las respuestas a las preguntas presentadas.

Líneas de Investigación, Desarrollo e Innovación

Esta línea de investigación contempla una serie de aportaciones al proceso de implantación de sistemas informáticos en el contexto de la industria nacional. Dentro de estos aportes, se puede mencionar la evolución del **Mo**delo de **Pro**ceso de **IMP**lantación de sistemas informáticos, denominado MoProIMP; su primera aproximación ha sido presentada en el Congreso Argentino de Ciencias de la Computación (CACIC 2018) (Panizzi M. et al., 2018).

En este contexto, se realizará una propuesta de métricas para el proceso en estudio, estas contemplarán una visión tridimensional anteriormente presentada. En concordancia a la visión tridimensional del proceso en estudio, se diseñarán un conjunto de recomendaciones riesgos como así también los procedimientos para su mitigación.

Resultados y Objetivos

En este estadío de la línea de investigación, los resultados logrados son:

- La experimentación del uso métodos de investigación en Ingeniería de Softtware Basada en Evidencia, el SMS y métodos de evaluación, DESMET.
- La construcción del establecimiento del estado del arte de la pregunta de investigación PI1 (Panizzi M. et al., 2019).
- La investigación documental del estado del arte de las respuestas de investigación PI2 y PI3 (Vázquez P., 2019), (Ortiz F., 2019).
- Una aproximación del modelo de implantación de sistemas informáticos y una validación inicial del mismo (Panizzi M. et al., 2018).
- Una propuesta de clasificación de métricas (Vázquez P. et al, 2018).
- La construcción y una validación inicial de una métrica de estimación para el proceso en estudio (Vázquez P. et al, 2019).

Los futuros pasos a seguir:

- Refinar el modelo de implantación y experimentar en diferentes casos de estudio de la industria nacional.
- Continuar con la construcción de las métricas como su validación en la industria nacional.
- Diseñar el conjunto de riesgos y sus procedimientos de mitigación como asi también sus respectivas validaciones en la industria.

Formación de Recursos Humanos

El equipo de investigación se encuentra conformado por un Director, dos investigadores formados, un investigador de apoyo, un investigador-estudiante, tres alumnos de posgrado. Se pretenden realizar dos tesis de Maestría y una tesis de Doctorado.

Referencias

Panizzi, M., Bertone R., Hossian A. (2017). Proceso de Implantación de Sistemas Informáticos – Identificación de vacancias en Metodologías Usuales. Libro de Actas de la V Conferencia Iberoamericana de Computación Aplicada CIACA 2017. Pag.207 -215. ISBN 978-989-8533-70-8. Vilamoura, Algarve, Portugal.

Kitchenham, B. and Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Versión 2.3 EBSE-2007-01.

Genero Bocco Marcela, Cruz-Lemus Jose Antonio y Piattini Velthuis Mario. (2014). Métodos de investigación en ingeniería del software. Editorial Ra-Ma, 2014. ISBN 978-84-9964-507-0.

Kitchenham, B. A. (1996). "DESMET: A Method for Evaluating Software Engineering Methods and Tools," Department of Computer Science, University of Keele, Technical Report TR96-09, August 1996. ISSN:1353-7776.

Panizzi, M., Hossian, A., García-Martínez, R. (2016). Implantación de Sistemas: Estudio Comparativo e Identificación de Vacancias en

Metodologías Usuales. Libro de Actas del XXII Congreso Argentino de Ciencias de la Computación. Universidad Nacional de San Luis. Pág. 546-555. ISBN 978-987-733-072-4.

Vázquez P., Panizzi M., Bertone R. (2018). Estimación del esfuerzo del proceso de implantación de software basada en el método de puntos de caso de uso. Mar del Plata, 29 y 30 de Noviembre de 2018 Universidad Tecnológica Nacional. 6to. Congreso Nacional de Ingeniería Informática / Sistemas de Información (CoNaIISI 2018). Simposio de Ingeniería de Sistemas y de Software. (En línea) ISSN: 2347-0372.

Panizzi, M., Bertone, R. Hossian, A. (2018). Propuesta de un Modelo de Proceso de Implantación de Sistemas Informáticos (MoProIMP). Libro de Actas del XXIV Congreso Argentino de Ciencias de la Computación. Universidad Nacional del Centro de la Provincia de Buenos Aires. ISBN: 978-950-658-472-6

Panizzi M. (2019). Establecimiento del estado sobre el proceso de implantación de sistemas informáticos. Reporte Técnico de Tesis Doctoral. RT-UNLP- III-LIDI. Publicado en: http://sedici.unlp.edu.ar/handle/10915/72633

Vázquez P. (2019). Estado del arte sobre métricas para el implantación de sistemas informáticos. Reporte Técnico del Trabajo de Especialización en Ingeniería en Sistemas de Información. Escuela de Posgrado. UTN-FRBA. Publicado en: ria.utn.ar/ handle/123456789/1841

Ortiz F. (2019). Establecimiento del estado sobre la gestión de riesgos en el proceso de implantación de sistemas informáticos. Reporte Técnico del Trabajo de Especialización en Ingeniería en Sistemas de Información. Escuela de Posgrado. UTN-FRBA. Publicado en: ria.utn.ar/ handle/123456789/1841

Pablo Vázquez; Marisa Panizzi, Rodolfo Bertone. (2019). Estimación del esfuerzo del proceso de implantación de software basada en el método de Puntos de Caso de Uso / Estimating the effort of the software implantation process based on the Use Case Points method. Brazilian Journal of Development, Curitiba, v. 5, n. 2, feb. 2019. Págs. 1809-1822. ISSN: 2525-8761.
http://www.brjd.com.br/index.php/BRJD

Aporte de los modelos semánticos a los sistemas de información

Romagnano, M.[1,2], Gómez, M. C.[1], Orellana, A.[1], Aballay, A.[1], Becerra, M.[1], Sarmiento, A.[1], Pantano, J.[1,2], Parra, L.[1], García, A.[1], Navarro, M.[1]

[1]Departamento de Informática, FCEFN, Universidad Nacional de San Juan
[2]Instituto de Informática, FCEFN, Universidad Nacional de San Juan

maritaroma@iinfo.unsj.edu.ar, cacugomez@yahoo.om.ar, arellana@iinfo.unsj.edu.ar, prof.alicia@gmail.com, mcbecerra2008@gmail.com, adriva2005@yahoo.com.ar, juancruz871@hotmail.com, garcia.exe@gmail.com, mirtaenavarro@yahoo.com.ar

RESUMEN

Desde sus orígenes y a través de su evolución los Sistemas de Información han tenido un rol protagónico en la sociedad. Así, estos sistemas impactan profundamente, y a nivel global, en el accionar continuo de organizaciones, empresas y/o instituciones; públicas o privadas. En la mayoría de los casos se cuenta con más de un Sistemas de Información para gestionar tareas centrales. Sin embargo, si bien continuamente se producen avances tecnológicos en cuanto a plataformas, movilidad, software, hardware, interacción sistemas-personas, entre otros; los sistemas siguen presentando problemas de interoperabilidad e inconsistencia. Además, la cantidad de usuarios que manipula y realiza transacciones con los Sistemas de Información crece exponencialmente a diario.

Aunque existen varios trabajos que abordan la problemática planteada, actualmente los Sistemas de Información siguen manteniendo el inconveniente de no poder ser aprovechados en su máximo potencial; contando con un abanico de conceptualizaciones.

Por lo tanto, debido a la importancia de los Sistemas de Información en cualquier ámbito social y puesto a la dificultad que aún mantienen, se propone asociar a éstos con modelos semánticos que aporten un patrón conceptual de datos, permitiendo el intercambio de información a través de la semántica de las instancias, sin necesidad de conocer los metadatos.

***Palabras Clave*:** Conceptual, Información, Modelos, Sistemas, Semántica.

CONTEXTO

El presente trabajo se encuentra enmarcado en el proyecto denominado "Aplicación de Modelos Semánticos a los Sistemas de Información (SI) para la mejora en el uso e impacto de las Tecnologías de Información (TI)", presentado en la convocatoria del Consejo de Investigaciones Científicas y Técnicas y de Creación Artística (CICITCA) de la Universidad Nacional de San Juan, para ser desarrollado durante el período comprendido entre 01/01/2018 al 31/12/2019, Código: 21/E1084.

Las tareas de investigación se desarrollan en el ámbito de la Facultad de Ciencias Exactas, Físicas y Naturales, con el apoyo del Departamento e Instituto de Informática.

El grupo de investigación, en su mayoría, tiene una experiencia de 20 años en la disciplina. Desde el año 1995 hasta el 2010 la línea de investigación estuvo orientada a la gestión de los SI. Luego, en el año 2011 se puso énfasis en los fundamentos conceptuales de los SI, pudiendo identificar las restricciones que existen en los SI. Se continuó estudiando la crisis por la que transitaba la disciplina al confundir los SI con TI. Posteriormente se trabajó en la construcción de un modelo disciplinar, basado en constructos. Desde el año 2016 se vienen identificando y caracterizando los modelos conceptuales que aportaran conocimiento a los SI.

En resumen, el actual proyecto cuenta con seis proyectos anteriores como precedentes.

1. INTRODUCCIÓN

Los SI cumplen un rol fundamental en la vida organizacional proveyendo de recursos esenciales, donde la información debe tener las cualidades necesarias para que los errores, las desviaciones y los problemas se solucionen y provean la mejora de la organización como sistema, en pos de la excelencia.

Al mismo tiempo, la globalización, la internacionalización de las organizaciones, el incremento de la competencia en los mercados de bienes y servicios, la rapidez en el desarrollo de las TI, el aumento de la incertidumbre en el entorno y la reducción de los ciclos de vida de los productos originan que la información se convierta en un elemento clave para la gestión, así como para la supervivencia y crecimiento de la organización empresarial (Hernandez Trasobares, 2003).

Si bien se han divulgado varias definiciones de SI, una de las más populares es la propuesta por Laudon y Laudon: "Sistema de Información es aquel conjunto de componentes interrelacionados que capturan, almacenan, procesan y distribuyen la información para apoyar la toma de decisiones, el control, análisis y visión de una organización" (Laudon y Laudon, 1996).

Todo SI utiliza como materia prima los datos, los cuales almacena, procesa y transforma para obtener como resultado final información, la cual será suministrada a los diferentes usuarios del sistema, existiendo además un proceso de feedback, en la cual se ha de valorar si la información obtenida se adecua a lo esperado. Pero, además, para poder realizar esta tarea un SI puede requerir de datos provistos por otro u otros sistemas. Estas transacciones pueden llevar a problemas de interpretación y/o interoperabilidad entre ellos.

Consecuentemente, resulta necesario que ellos se encuentren apoyados por un modelo conceptual que proporcione una explicación detallada de cada uno de estos.

En un SI, donde probablemente existe una vasta cantidad de información por manipular, se hace necesario tener presente la semántica y aún, si se requiere más precisión, la sintaxis. Una forma de abordar la primera es trabajar con modelos semánticos y/o bases de datos léxicas que sirvan de soporte a este sistema. Una base de datos léxica y morfológica proporciona un acabo conjunto de vocablos o términos que enriquecen el conocimiento de un contexto. Generalmente un vocabulario léxico representa la forma de expresión de la lengua en un determinado grupo social. Según la Real Academia Española, léxico se refiere al conjunto de palabras de un idioma, o de las que pertenecen al uso de una región, a una actividad determinada, a un campo semántico dado, etc. La morfología explica la estructura interna de las palabras y el proceso de formación de palabras.

La semántica es la parte de la lingüística que estudia el significado de las expresiones lingüísticas y la sintaxis describe cómo las palabras se combinan para formar sintagmas, oraciones y frases (Princeton University, 2016).

Por lo tanto el léxico se relaciona con la semántica y la morfología con la sintaxis.

Los modelos semánticos comúnmente usados para contextualizar una realidad se simplifican en taxonomías, tesauros y ontologías.

Pérez Porto y Merino definieron a la taxonomía como la ciencia que estudia los principios, métodos y fines de la clasificación (Pérez Porto y Merino, 2013).

Un tesauro se refiere a un glosario estructurado. Puede ser un listado que presenta los términos que se utilizan para la representación de ciertos contenidos. Su finalidad es contribuir a la normalización de los términos; de este modo, se facilita la comprensión y la comunicación. De esta manera, en los tesauros se puede hallar un listado de los términos preferidos (ordenados de modo jerárquico, temático y alfabético), un listado de descriptores (sinónimos de los términos preferidos), las relaciones existentes entre dichos términos y, por último, las definiciones de estos términos. Además los tesauros suelen incluir reglas para su uso (Pérez Porto y Gardey, 2014).

Según Castell, una ontología es una taxonomía de conceptos con atributos y relaciones, que proporciona un vocabulario consensuado para definir redes semánticas de unidades de información interrelacionadas (Castell, 2002). Por su parte, Nykänen define que una ontología es una especificación formal y explicita de una conceptualización compartida de un dominio particular. Para lograr esto, la ontología debe estar conformada por una taxonomía y un conjunto de reglas de inferencia (Nykänen, 2003).

En el contexto de las Ciencias de la Información y la Computación, una ontología se define como conjunto de primitivas representacionales con lo cual se modela un dominio de conocimiento o discurso. Estas, típicamente son clases, atributos y relaciones (o relaciones entre los miembros de las clases) e incluyen información sobre significado y sus limitaciones sobre su aplicación lógicamente coherente. Las ontologías están generalmente especificadas en lenguajes que permiten la abstracción fuera de las estrategias de implementación y de las estructuras de datos. Por esta razón, se dice que las ontologías están en un nivel "semántico", mientras que los esquemas de base de datos son modelos de datos del nivel "físico" o "lógico". Debido a su independencia de los modelos de datos de menor nivel, las ontologías se usan para integrar bases de datos heterogéneas, con lo cual se logra la interoperabilidad entre sistemas dispares. Así, los programas informáticos pueden utilizar las ontologías con una variedad de propósitos, incluyendo el razonamiento inductivo, la clasificación y una variedad de técnicas de resolución de problemas (Gruber, 2009).

Ya en el año 2007 Barchini, Álvarez, Herrera y Trejo habían señalado que el uso de ontologías en el desarrollo de los SI permitiría establecer correspondencia y relaciones entre los diferentes dominios de entidades de información (Barchini, Álvarez, Herrera y Trejo, 2007).

Además, Codina y Jiménez-Pedraza realizaron una comparación entre tesauros y ontologías para tratar el tema de la semántica relacionada con SI documentales. Plantearon que un tesauro es una tecnología sólida en los SI, con modelos bien establecidos y con una buena implantación en la industria, mientras que la ontología es una tecnología prometedora y presenta un enorme potencial, ya probada en otros campos (por ejemplo, en la lingüística y en ámbitos específicos de la inteligencia artificial). También insinuaron que incorporar ontologías a los SI proporcionaría las bases para la interconexión de datos y la prestación de mejores servicios de información. En este trabajo los autores mencionaron: "gracias a las ontologías y al proyecto de la web semántica, estamos de nuevo en una situación en la que es imposible enfocar el telescopio sin descubrir alguna cosa" (Codina y Jiménez-Pedraza, 2011).

En cuanto a su aplicación, hoy en día los SI están presentes en distintos estratos sociales. Por ejemplo, en el trabajo de Dorado, González Serrano y Jiménez Builes se planteó que existen varios sistemas de gestión de aprendizaje ocasionando un problema de semántica en el conocimiento, haciéndolo disperso e inexacto debido a que cada uno de estos sistemas maneja sus propios formatos, arquitecturas y especificaciones. Aquí es donde se devela claramente la carencia de interoperabilidad de la información entre diferentes plataformas, afectando la efectiva gestión de las instituciones. Los autores propusieron el uso de ontologías aplicables a e-learning para desarrollar el modelo (Dorado, González Serrano y Jiménez Builes, 2015).

Un caso de estudio concreto que se puede mencionar es el de la cadena multinacional de supermercados minoristas "Walmart". Una de las claves de su éxito fue la implementación de Retail Link (Sidebar), un sistema de gestión de suministros. Este sistema, único cuando se implementó inicialmente a mediados de la década de 1980, permitió a los proveedores de Walmart acceder directamente a los niveles de inventario e información de ventas de sus productos. Al utilizar Retail Link, los proveedores pueden analizar qué tan bien se venden sus productos en cualquiera de las sucursales, con una gama de opciones de

informes. Además, Walmart requiere que los proveedores utilicen Retail Link para administrar sus propios niveles de inventario. Si un proveedor considera que sus productos se están vendiendo demasiado rápido, pueden usar este sistema para solicitar al supermercado que aumente los niveles de inventario. Básicamente, esto ha permitido a Walmart "contratar" a miles de gerentes de productos, los cuales tienen un gran interés en los productos que están administrando y continuar reduciendo los precios, así como responder rápidamente a las fuerzas del mercado (Bourgeois y Bourgeois, 2014).

Otro ejemplo de aplicación de un SI es desde el punto de vista de la Web, en cuyo caso se lo contempla como un buscador. Desde esta mirada, tanto la Web 2.0, como la 3.0 y la Web Semántica han tenido impactos moderados en el núcleo funcional de los SI generalistas en la web. Así, los entornos intensivos en información son uno de los sectores que más están aprovechando estas nuevas tecnologías de la Web, tanto en forma de sistemas de gestión de información personal, como en nuevas prestaciones de las bases de datos. La abundancia de soluciones ante una consulta, sistemas que se solapan parcialmente, fuentes de información cruzadas, alternativas en el uso de recursos informativos, etc., hace que, en el panorama actual sea por un lado más fácil que nunca encontrar información, pero más complicado organizarla de forma eficaz (Codina, 2009).

Resumiendo y teniendo en cuenta los antecedentes planteados, se identifica la necesidad de contar con un modelo semántico que apoye el desempeño de un SI, cualquiera sea la institución u organización en el cual esté inmerso.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

El proyecto presenta características interdisciplinarias, debido a que se consideran como pilares fundamentales los datos, los recursos humanos y la tecnología; en los que se sustentan los SI.

Los datos representan el “corazón” de los SI, los recursos humanos son quienes manipulan dichos sistemas y a quienes benefician o no. Las tecnologías proveen los medios para el funcionamiento automático de dichos sistemas.

Por lo tanto, actualmente el grupo se encuentra realizando tareas divididas de la siguiente forma:

- *Modelos semánticos*. Se realizó un estudio y análisis de antecedentes y de la bibliografía existente.
- *Modelos semánticos existentes y que actualmente se aplican a los SI*. Se ha realizado un estudio y análisis de antecedentes. Actualmente se está modelando y desarrollando una ontología.
- *Implicancias organizacionales y legales al aplicar modelos semánticos a los SI*. Se analizarán las compatibilidades organizacionales y legales de incorporar modelos semánticos a los SI.

3. OBJETIVOS ESPERADOS

Como objetivo general se espera identificar y proponer modelos semánticos que aporten significancia a los SI.

Como objetivos específicos se han propuesto:

- Estudio y análisis de material bibliográfico, estado del arte y antecedentes relacionados con la problemática.
- Aplicar técnicas y herramientas proporcionadas por las Ciencias del Comportamiento, de la Semántica, de la Ingeniería de Software y de la Inteligencia Artificial; para conceptualizar los SI.
- Obtener conclusiones sobre los aportes de los modelos conceptuales a los SI.

4. FORMACIÓN DE RECURSOS HUMANOS

Actualmente el proyecto cuenta con un alumno de doctorado; en etapa de revisión y defensa de tesis, un alumno de maestría y uno de la Licenciatura en SI; en etapa de escritura de tesis. A fines del año 2018 se ha licenciado un alumno desarrollando su tesis en la temática del proyecto y actualmente continúa formándose como investigador en SI.

Debido a que dos investigadores (pioneros y referentes en SI) se han jubilado, se prevé la incorporación de profesionales en la disciplina.

5. BIBLIOGRAFÍA

Barchini, G., Álvarez, M., Herrera, S., y Trejo, M. (2007). "El Rol de las Ontologías en los SI". Revista Ingeniería Informática, Ed. 14. Disponible en: http://www.inf.udec.cl/revista

Bourgeois, D. y Bourgeois, D. (2014). "Information Systems for Business and Beyond. A look at the technology, people, and processes of information systems". Disponible en: https://bus206.pressbooks.com/

Castell, P. (2002). "Aplicación de técnicas de la web semántica". Escuela Politécnica Superior. Universidad Autónoma de Madrid. Disponible en: http://arantxa.ii.uam.es/~castells/publications/coline02.pdf

Codina, L. y Jiménez-Pedraza, R. (2011). "Tesauros y Ontologías en SI Documental". El profesional de la información, septiembre-octubre, V. 20, Nro. 5. Disponible en: http://www.elprofesionaldelainformacion.com/contenidos/2011/septiembre/10.pdf

Codina, L. (2009). "¿Web 2.0, Web 3.0 o Web Semántica?: El impacto en los SI de la Web". I Congreso Internacional de Ciberperiodismo y Web 2.0. Bilbao. Disponible en: https://www.lluiscodina.com/wp-content/uploads/Web20_WebSemantica2009_Nov2009.pdf

Dorado, A., González Serrano, C., y Jiménez Builes, J. (2015). "Model of Semantic interoperability between Learning Management Systems". Disponible en: http://revistas.sena.edu.co/index.php/inf_tec/article/view/137/370

Gruber, T. (2009) "Ontology. Encyclopedia of Database Systems", Ling Liu and M. Tamer Özsu (Eds.), Springer-Verlag, 2009. Disponible en: http://tomgruber.org/writing/ontology-definition-2007.htm

Hernández Trasobares, A. (2003). "Los SI: Evolución y Desarrollo". Proyecto social: Revista de relaciones laborales, ISSN 1133-3189, Nº 10-11, 2003, págs. 149-165. Disponible en: https://dialnet.unirioja.es/servlet/articulo?codigo=793097

Laudon, K. y Laudon, J. (1996). "Administración de los SI", Prentice Hall, México.

Nykänen, O. (2003). "Metadata for Learning Resources: Technologies and Directions of the Semantic Web – A Brief Review". IEEE International Conference on Advanced Learning Technologies (ICALT'04).

Pérez Porto, J. y Merino, M. (2013). "Definición de Taxonomía". Disponible en: https://definicion.de/taxonomia/

Pérez Porto, J. y Gardey, A. (2016). "Definición de Tesauro". Disponible en: https://definicion.de/tesauro/

Princeton University, (2016). "Princeton University. WordNet". Disponible en: https://wordnet.princeton.edu/wordnet/related-projects/#semantic.

ARQUITECTURA INFORMÁTICA PARA LA IMPLEMENTACIÓN DE LA GESTIÓN DEL CONOCIMIENTO EN UNA CÁTEDRA UNIVERSITARIA

Straccia, Luciano[1]; Saavedra Martínez, Paola[1,2]; Zángara, Alejandra[2] ; Pollo-Cattaneo, Ma. Florencia[1]

[1]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

[2]Programa de Maestría en Tecnología Informática Aplicada en Educación
Universidad Nacional de La Plata, Facultad de Informática.
Calle 50 y Calle 120 (B1900ASD) La Plata. Buenos Aires. Tel +54 221 427-7270

lstraccia@frba.utn.edu.ar, psaavedra_martinez@hotmail.com, alejandra.zangara@gmail.com, flo.pollo@gmail.com

Resumen

El conocimiento constituye un recurso estratégico y su gestión otorga ventajas competitivas. Existen diversos modelos que han sido creados para guiar la Gestión del Conocimiento en una organización, pero es necesario trabajar específicamente sobre modelos de arquitectura de tecnología de la información que favorezcan su aplicación. En este contexto, el objetivo del proyecto es definir una arquitectura informática y su metodología de aplicación en la gestión de conocimiento en cátedras universitarias.

Palabras clave: Gestión del Conocimiento, gestión universitaria, arquitectura tecnológica, Universidad, Ingeniería de Software.

Contexto

La Universidad Tecnológica Nacional (UTN) ha definido diversos Programas de Investigación + Desarrollo + Innovación (I+D+i), entre los cuales se encuentra el Programa Tecnología de las Organizaciones, que tiene por finalidad "articular las distintas temáticas relacionadas con la gestión del conocimiento, la innovación y los sistemas de gestión de la calidad e integrados aplicables a las organizaciones", incluyendo la gestión del conocimiento en las organizaciones como área prioritaria [1], bajo la cual se llevan adelante las actividades detalladas en este trabajo.

Introducción

<u>Conocimiento y gestión del conocimiento</u>

La Gestión del conocimiento (GC) es un proceso a través del cual las organizaciones logran descubrir, utilizar y mantener el conocimiento, con la idea de alinearlo con las estrategias de negocio para la obtención de ventajas competitivas [2]. Este proceso debe proveer el entorno tecnológico para

facilitar dicha gestión. El conocimiento "se constituye en el recurso estratégico más importante, y la habilidad para generarlo, adquirirlo, codificarlo, transferirlo, aplicarlo y reutilizarlo, se ha convertido en la competencia sustancial para la obtención de una ventaja competitiva sostenible" [3].

En lo que respecta a la GC, existen numerosas publicaciones y la mayoría se basa en el modelo de gestión de Nonaka y Takeuchi [4], quienes introducen la idea de la conversión del conocimiento tácito en explicito, definiendo al tácito como aquel que es difícil de expresar, formalizar y compartir, muy personal y subjetivo, derivado de la experiencia; y al conocimiento explícito como aquel que se puede expresar y formalizar fácilmente, por lo cual, se adquiere mediante métodos formales de estudio. Diversos autores como Davenport [5], Davenport y Prusak [6], Wiig [7] y Pérez y Urbáez [8] han presentado otras definiciones asociadas al conocimiento.

GC en Universidades

Partiendo de conceptos de Fabre Batista [9], se pueden definir algunos de los objetivos principales que se deben lograr en la gestión de una cátedra universitaria: consolidar el proceso de crecimiento de la docencia, promover la creación de espacios académicos para el perfeccionamiento, formación de los docentes y la investigación y ejecución de actividades vinculadas a extensión.

Para lograr estos objetivos se debe trabajar colaborativamente de forma tal que se puedan compartir tanto los procesos de generación como el conocimiento mismo. Esta colaboración debería involucrar el intercambio de producción teórica, metodologías y técnicas dentro de la Universidad.

Algunos interrogantes fundamentales que se presentan en este contexto son ¿cómo la tecnología informática colabora con los docentes para llevar estas actividades de intercambio de conocimiento dentro de la cátedra?, ¿cómo visibilizar, a través de la tecnología, las distintas maneras en que equipo de cátedra comparte el conocimiento y se relaciona con la investigación y la extensión?

Entre los aspectos que la gestión de cátedras debe considerar para cumplir sus objetivos, puede hallarse la integración de docentes nuevos a las ideas de la cátedra; vinculación con otras materias relacionadas; homogeneidad en los contenidos: si bien cada docente incorpora su propio know-how, es fundamental contar con una base de conocimiento (entendida como el conjunto de conocimiento, ideas, conceptos independientemente de su implementación).

Respecto a GC en universidades, en [10] se resumen teorías y experiencias en universidades y se presenta una teoría del conocimiento que reconoce la interacción de la realidad, de los sujetos, del sistema de ideas y creencias, del lenguaje y del consenso social. En [11] los autores se centran en la necesidad de las organizaciones educativas respecto de nuevas formas de organización que sean capaces de poner en práctica nuevas soluciones basadas en la gestión del conocimiento. En [12] se aborda el significado de la gestión del conocimiento en la transformación institucional de los planteles del sistema educativo escolar particularmente en México. En [13] se pone el foco en cómo identificar, medir y evaluar los activos componentes del Capital Intelectual y qué directrices o programas se pueden formular para orientar la dirección y gestión del conocimiento. En tanto, en [14] se

presenta la aplicación del modelo de gestión de Nonaka y Takeuchi en el marco universitario y en [15] se presenta una revisión documental acerca de la gestión del conocimiento y su situación actual en una universidad venezolana.

TICs para la gestión del conocimiento

En lo referido a las Tecnologías de la Información y Comunicación (TICs) para la gestión del conocimiento, se parte de preguntas como: ¿qué tipo de TIC puede contribuir al flujo de conocimiento, apoyando su conversión desde explícito a tácito y, desde tácito a explícito?, ¿qué tipos de TIC pueden apoyar mejor al conocimiento explícito que una organización tiene?, ¿cómo gestionar, a través de las TIC, el volumen de conocimiento explícito contenido en las colecciones de documentos de una organización?

Diversos autores han mencionado algunas de estas tecnologías o herramientas en sus modelos de gestión del conocimiento. Borghoff et al. [16] proponen un modelo sociotécnico (con uso de tecnología informática) basado en el modelo de Nonaka y Takeuchi y persiguen conocer las tecnologías necesarias para facilitar la GC. El modelo se centra en estudiar las tecnologías de la información y cómo pueden contribuir al flujo de conocimiento, apoyando su conversión desde explícito a tácito y desde tácito a explícito. La memoria corporativa registra el conocimiento acumulado sobre los servicios y los productos de una organización, con el propósito de apoyar la mejora continua de prácticas de trabajo intensivas en conocimiento. El modelo propone una estructura de cuatro componentes: depósitos y almacenes de conocimiento; comunidades de trabajadores del conocimiento; flujo del conocimiento; cartografía del conocimiento.

Peter Tyndale [17] distingue entre tecnologías de la información basadas en herramientas tomadas prestadas de otras disciplinas que han entrado en el campo de la GC, y tecnologías de la información basadas en herramientas específicas. Ambos tipos de herramientas pueden estar, y de hecho se usan, dentro de la gestión del conocimiento. Las herramientas definidas son: intranet, portales web, gestión de contenidos, motores de recuperación, sistemas de gestión electrónica documental, base de datos relacionales, almacenes de datos, aplicaciones de mesa de ayuda, reingeniería de procesos, agentes de software inteligentes y aplicaciones de creación de conocimiento.

Carvalho et al. [18] desarrollan una tipología de soluciones que comprende diez categorías: sistemas basados en intranets, gestión de documentos electrónicos, groupware, workflow, sistemas basados en conocimiento, inteligencia de negocios, mapas de conocimiento, herramientas de apoyo a la innovación, herramientas de inteligencia competitiva y portales de conocimiento.

Por otra parte, algunos de los modelos definen particularmente el uso de las TICs, pero no como un modelo integral. Wiig [19] se enfoca en que un conocimiento puede ser útil y valioso solo si puede ser organizado y define que los procesos básicos para la GC deben ser: creación, captura, renovación, distribución y uso del conocimiento. En este modelo se visualiza la gestión del conocimiento proyectado hacia la organización en general, el aprendizaje formal y el conocimiento explícitamente documentado. Además, solo se destaca el uso de las Tics en la distribución del conocimiento utilizando sistemas basados en conocimiento y redes de expertos.

Por otro lado, se visualiza una falencia metodológica para la implementación de aspectos técnicos en la GC; se mencionan y detallan tecnologías disponibles, pero no existe arquitectura o modelo que defina particularmente el uso de las variadas tecnologías existentes, como una arquitectura integrada. Lo más parecido a una arquitectura completa es el modelo de Kerschberg [20] que propone un modelo de 3 capas donde se podrían utilizar algunas herramientas tecnológicas en alguna de esas capas, pero no explicita claramente la implementación de TIC en el marco de la arquitectura presentada.

Si se analizan las tecnologías detalladas por los primeros autores mencionados se puede determinar que se podría generar un modelo, basado en el de Kerschberg, que permita concretamente definir un camino de "buenas prácticas" para lograr la aplicación de tecnologías eficientemente en la GC en una cátedra. Es en este aspecto donde se encuentra una brecha entre los modelos existentes de GC y su implementación tecnológica.

Líneas de Investigación, Desarrollo e Innovación

En la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN.BA) se ha conformado en el año 2009 el Grupo GEMIS, con dependencia del Departamento de Ingeniería en Sistemas de Información, integrado por un equipo de docentes, alumnos y graduados con interés en la sistematización de conocimientos y su promoción sobre el campo de la Ingeniería en Sistemas de Información y la Ingeniería de Software incluyendo sus aplicaciones y abordajes metodológicos en todo tipo de escenarios. Se ha iniciado un nuevo Proyecto de Investigación y Desarrollo (PID) denominado "La gestión del conocimiento en pequeñas y medianas fábricas de software en el Área Metropolitana de Buenos Aires" que indaga en los modelos de gestión del conocimiento. A partir de las experiencias y resultados obtenidos en dicho trabajo se comienza una nueva línea de trabajo vinculada específicamente a la arquitectura tecnológica, sus herramientas y sus tecnologías. En primer lugar se analizará en el marco de la gestión de cátedras universitarias, para luego ampliar los resultados hacia las diferentes organizaciones de Ingeniería de Software y fábricas de software. De esta manera, se articula dentro de los objetivos de GEMIS en el campo de la Informática, la generación de nuevos conocimientos en el área de la Ingeniería de Software.

Resultados esperados

Objetivo General

Definir una arquitectura informática y su metodología de aplicación en la gestión de conocimiento en cátedras universitarias.

Objetivos Específicos

Los objetivos específicos son: a) identificar los principales desafíos que presenta la gestión del conocimiento en la gestión de una cátedra universitaria; b) identificar modelos de gestión del conocimiento adecuados a los desafíos que presenta la gestión del conocimiento en la gestión de una cátedra universitaria; c) proponer una arquitectura de implementación de tecnologías informáticas para la gestión del conocimiento en una cátedra universitaria; d) validar la arquitectura propuesta; e) exponer las conclusiones sobre el tema y posibles aportes futuros.

Resultados esperados

Se esperan obtener los siguientes resultados: a) definir una arquitectura informática para la gestión del conocimiento en una cátedra universitaria; b) definir de una metodología de implementación de la arquitectura informática para la gestión del conocimiento en dicho contexto; c) validar la metodología propuesta mediante el uso de la técnica de juicio de expertos. Esta técnica se define como una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencia, juicios y valoraciones [21].

Formación de Recursos Humanos

El equipo se encuentra conformado por investigadores formados, tesistas de maestría, graduados de grado y alumnos de la carrera de Ingeniería en Sistemas de Información. Esta nueva línea de trabajo busca tanto la obtención de nuevos conocimientos como la motivación de los implicados para que asciendan dentro de la carrera de investigadores, además de fomentar la aplicación de las temáticas en su propia actividad profesional; y plantea la integración de alumnos avanzados de grado y posgrado con posibilidades de articular Proyectos Finales, Trabajos Finales Integrador de Especialidad y Tesis de Maestría de las carreras del área.

Referencias

[1] SeCTyP (2017). Programas I+D+I. UTN. Secretaría de Ciencia, Tecnología y Posgrado.

[2] Bueno, E. (2000). Dirección del Conocimiento y Aprendizaje: Creación, distribución y mediación de Intangibles.

[3] Gelaf, Graciela (2010). Abordajes creativos en situaciones de crisis organizacionales. Contaduría General de la Nación. Tucumán, Argentina.

[4] Nonaka, I., & Takeuchi, H. (1995). The knowledge-creating company: How Japanese companies create the dynamics of innovation. Oxford university press

[5] Davenport, T. (1998). Knowledge Management and the broader fim: Strategy, advantage, and performance. En J. Liebowitz, Knowledge Management Handbook.

[6] Davenport, T. y Prusak, L. (1998). Working Knowledge: How organizations manage what they know. Cambridge, MA: Harvard Business School Press.

[7] Wiig, K. (1993). Knowledge Management Foundations: Thinking about thinking – How people and organizations create, represent, and use knowledge.

[8] Pérez, V. A., & Urbáez, M. F. (2016). Modelos teóricos de gestión del conocimiento: descriptores, conceptualizaciones y enfoques. Entreciencias: diálogos en la Sociedad del Conocimiento, 4(10), 201-227.

[9] Fabre Batista, G (2009). Las funciones sustantivas de la universidad y su articulación en un departamento docente. Universidad Agraria de La Habana.

[10] Perez Lindo, A (2017). El uso social del conocimiento y la universidad. 1a ed. Universidad Abierta Interamericana, 2017.

[11] Sosa, M; Velázquez, I; Zarco, R; Acevedo, J (2010). Modelo de Gestión Académica para Instituciones Educativas Basado en el enfoque de Gestión del Conocimiento. Universidad Nacional de Santiago del Estero. WICC 2010.

[12] Minakata Arceo, A. (2009). Gestión del conocimiento en educación y transformación de la escuela. Notas para un campo en construcción. Sinéctica, Revista Electrónica de Educación, (32), 1-21.

[13] Morcillo Ortega, J y otros. (2003) Dirección General de Investigación, Consejería de Educación. Comunidad de Madrid. M-54391-2003

[14] García Cienfuegos, B (2016). Gestión del Conocimiento desde el contexto universitario.

[15] González de Becerra, Y; Frassati, E. (2010). Gestión del Conocimiento en el Área de Investigación de las Universidades Públicas. Revista electrónica de Humanidades, Educación y Comunicación Social.

[16] Borghoff, U y Parechi, R (1998) Information Technology for Knowledge Management

[17] Tyndale, Peter (2002). A taxonomy of knowledge management tools: origins and applications. Evaluation and Program Planning 25 (183-190).

[18] Carvalho y Ferreira (2001). A Proposal of Taxonomy for Knowledge Management Tools.

[19] Wiig, K. (1993). Knowledge Management Foundations: Thinking about thinking – How people and organizations create, represent, and use knowledge.

[20] Kerschberg, L y Weishar, J (2000). Conceptual Models and Architectures for Advanced Information Systems

[21] Escobar J, Cuervo A (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. Avances en medición, 6(1), 27-36.

ASPECTOS DE INGENIERIA DE SOFTWARE Y BASES DE DATOS PARA EL DESARROLLO DE SISTEMAS DE SOFTWARE EN ESCENARIOS HIBRIDOS.

Luciano Marrero, Pablo Thomas, Ariel Pasini, Rodolfo Bertone, Eduardo Ibáñez, Alejandra Rípodas, Verónica Aguirre, Verena Olsowy, Fernando Tesone, Magalí Capecci, 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

{lmarrero, pthomas, apasini, pbertone, eibanez, aripodas, vaguirre, volsowy, ftesone, mcapecci, ppesado}@lidi.info.unlp.edu.ar

RESUMEN

Se presenta una línea de investigación que tiene por objeto estudiar el desarrollo de sistemas de software, la influencia de las bases de datos relacionales, el impacto de las bases de datos no relacionales y los procesos actuales de Ingeniería de Software en escenarios híbridos. Se tiene como objetivo principal generar metodologías y prácticas de Ingeniería de Software considerando las características de estos escenarios, así como también realizar un análisis cualitativo y cuantitativo de la información generada y los aspectos de usabilidad que caracterizan a un sistema de software que debe tener en cuenta los avances de la tecnología y el auge de uso en diferentes contextos. Esto implica, tener en cuenta aspectos que hace algunos años no eran considerados, tales como, movilidad, geolocalización, generación de grandes volúmenes de información y diversidad de dispositivos electrónicos involucrados.

Palabras claves: Ingeniería de Software, Bases de Datos, NoSQL, Escenarios Híbridos, Aplicaciones Móviles.

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 y Latinoamérica.

Se utilizan los recursos de Hardware y Software disponibles en el III-LIDI para diseñar, desarrollar y probar diferentes soluciones a problemáticas relacionadas con escenarios a investigar. Como resultado de esto, se espera obtener métricas reales que sirvan como referencia para los investigadores en la comparación de resultados.

1. INTRODUCCION

“Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso”, 2018-2021 del Programa de Incentivos, es un proyecto que propone profundizar las investigaciones que se vienen realizando en el III-LIDI y extender la mirada a nuevos desafíos y cambios que están en gestación. 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 [13].

Este artículo se centra en el subproyecto “SP1 -Ingeniería de Software para escenarios híbridos”. Se orienta a la investigación de metodologías y técnicas de la Ingeniería de Software y Bases de Datos, con énfasis en los escenarios híbridos, aplicaciones móviles y la necesidad de personalización y adaptación acorde al contexto y al estado de cada usuario.

La creciente disponibilidad y acceso a dispositivos móviles, la expansión del acceso a internet, el uso de la geolocalización, el auge de las redes sociales y el uso de otras tecnologías, han marcado la profundización de nuevos entramados entre mundo físico y virtual. Se presenta así, nuevos desafíos para los principios metodológicos de la Ingeniería de Software y las arquitecturas tradicionales utilizadas para la representación de la información, dando lugar al surgimiento de una nueva generación de sistemas que debe responder a tales exigencias [1, 2 y 9].

Estos escenarios son reconocidos como híbridos, en los cuales aumenta la posibilidad de acceso y procesamiento de la información digital, ya que no solo es posible acceder desde una computadora de escritorio, sino también desde una variedad de dispositivos electrónicos con diferentes características.

Los escenarios mencionados anteriormente son integrados en diferentes ámbitos, modificando el concepto que se tiene de ellos, por ejemplo, las ciudades se convierten en inteligentes, la enseñanza convencional se extiende a entornos virtuales, se generan nuevas posibilidades para las actividades cotidianas y, además, se llama a la investigación y reflexión en el ámbito de las Ciencias de la Computación [3, 8, y 10].

A causa de la digitalización de la información, los sistemas de software han aumentado de forma significativa el volumen de datos a procesar y almacenar. Además, deben responder eficazmente al aumento de usuarios que lo acceden de forma simultánea. Esto significa que la escalabilidad y el rendimiento se han convertido en auténticos retos para la estructuración y el almacenamiento de la información. De modo que las tecnologías de bases de datos más comúnmente utilizadas, las relacionales (SQL), pueden no ser suficientes.

En respuesta a esta problemática, como una alternativa a las bases de datos relacionales surgen las bases de datos no relacionales (NoSQL), donde la información almacenada no necesariamente requiere de una estructura determinada, es decir, que operan sin un esquema predefinido. Estas bases de datos se caracterizan por ser altamente escalables y, aunque la gran mayoría no implementan estrictamente las propiedades ACID (Atomicidad, Consistencia, Aislamiento y Durabilidad), poseen características que las

hacen adecuadas para la administración de grandes volúmenes de datos [6, 7 y 12].

El desarrollo de software para escenarios híbridos requiere la modificación de las prácticas tradicionales de la Ingeniería de Software: gestión de requerimientos, diseño e implementación de Bases de Datos, modelos de proceso, etapas de diseño e implementación de interfaces y planes de prueba, entre otras.

En el desarrollo actual de proyectos de software, es fundamental la comunicación, coordinación y sincronización del trabajo. La utilización de repositorios de información, por ejemplo GIT, permiten realizar un 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 [4, 5, 10 y 11].

Como se mencionó anteriormente, el exponencial aumento del uso de aplicaciones móviles que se ha visto en los últimos años impuso un nuevo paradigma para el desarrollo de sistemas.

Se deben tener en cuenta ciertos aspectos de los dispositivos móviles que pueden afectar el desempeño de tales aplicaciones, recursos de almacenamiento y procesamiento limitado, conectividad variable, fuente finita de energía, etc. Si bien estos dispositivos poseen su propia forma de almacenar información, en algunos casos, cuando el volumen de datos a administrar es grande, puede ser necesaria la interacción con otros dispositivos, dando lugar a las bases de datos orientadas a dispositivos móviles. Estas son bases de datos portables que poseen cierta independencia del servidor y pueden comunicarse desde cualquier punto remoto para sincronizar y centralizar la información.

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 [2, 3, 7 y 12].

Todas las particularidades previamente mencionadas conducen a nuevas prácticas y metodologías que promueven el crecimiento de la Ingeniería de Software como disciplina.

2. LINEAS DE INVESTIGACION Y DESARROLLO

- Metodologías y Técnicas de la Ingeniería de Software y su aplicación en el desarrollo de software para escenarios híbridos.
- Técnicas, atributos y métricas de usabilidad en sistemas para escenarios híbridos
- Metodologías para la interoperabilidad de aplicaciones móviles y sistemas web.
- Bases de datos relacionales.
- Bases de datos no relacionales.
- Bases de datos móviles.
- Repositorios GIT.

- Trabajo colaborativo en Proyectos de Software.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados esperados/obtenidos se pueden resumir en:

- Capacitación continua de los miembros de las líneas de investigación.
- Analizar, comparar y adaptar las nuevas metodologías y herramientas de la Ingeniería de Software para el desarrollo del software.
- Definición de técnicas, atributos y métricas para el análisis de usabilidad de sistemas en escenarios híbridos.
- Estudio y análisis de Bases de Datos no relacionales.
- Comparación de Bases de Datos relacionales con Bases de Datos no relacionales.
- Definición de procesos de Gestión de Incidencias utilizando repositorios GIT.
- Análisis de metodologías para la interoperabilidad de sistemas web y aplicaciones móviles.

Algunas de las transferencias realizadas por el III-LIDI relacionadas con este proyecto son:

- 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).
- Sistema de Software para la gestión del comedor del Albergue Universitario de la UNLP.
- Aplicación móvil para la comunidad de la Facultad de Informática de la UNLP con información sobre horarios de finales, planes de estudio, calendario académico, las clases en tiempo real y las últimas novedades.

4. FORMACION 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.

5. BIBLIOGRAFIA

1. Ingeniería de Software. Un Enfoque Práctico. Séptima Edición. Roger S. Pressman. McGraw-Hill Interamericana Editores, S.A. 2010. ISBN: 978-607-15-0314-5
2. Ingeniería de Software Teoría y Práctica. Pfleegger Shari Lawrence. Pearson / Prentice Hall. 2002. ISBN: 9789879460719
3. Ingeniería de Software. Novena Edición. Ian Sommerville. Addison Wesley / Pearson. 2011. ISBN: 978-607-32-0603-7
4. Ingeniería de Software Clásica y Orientada a Objetos. Sexta Edición. Stephen R. Schach. Mc Graw Hill Interamericana Editores S.A. 2006. ISBN: 970-10-5636-1

5. Administración de Proyectos. Guía para el Aprendizaje. Francisco Rivera Martínez, Gisel Hernández Chávez. Prentice Hall / Pearson. 2010. ISBN: 978-607-442-620-5
6. Utilización de NoSQL para resolución de problemas al trabajar con cantidades masivas de datos. Róttoli, Giovanni, López Nocera, Marcelo, Pollo Cattaneo María Florencia. 2015. XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015). http://sedici.unlp.edu.ar/handle/10915/45514
7. NoSQL A Brief Guide To The Emerging World of Polyglot Persistence. Pramod J. Sadalage y Martin Fowler. Pearson Education. 2013. ISBN: 978-0-321-82662-6
8. Metodologías, técnicas y herramientas de ingeniería de software en escenarios híbridos. Bertone Rodolfo, Thomas Pablo Javier. Pasini Ariel, Marrero Luciano, Ibañez Eduardo, Ripodas Alejandra, Aguirre Verónica, Olsowy Verena, Capecci, Eva Magalí, Tesone Fernando, Pesado, Patricia. 2018. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). ISBN 978-987-3619-27-4
9. Dispositivos móviles: desarrollo de aplicaciones y conectividad. Thomas Pablo Javier, Galdámez Nicolás, Delía Lisandro Nahuel Cristina Federico, Dapoto Sebastián, Pesado Patricia. 2014. XVI Workshop de Investigadores en Ciencias de la Computación (Ushuaia 2014). ISBN: 978-950-34-1084-4
10. Ingeniería de Software para Sistemas Distribuidos. Pesado Patricia, Bertone Rodolfo, Thomas Pablo Javier, Marrero Luciano, Pasini, Ariel, Delía Lisandro Nahuel, Galdámez Nicolás, Ibañez Eduardo, Estrebou César Armando, Ripodas Alejandra, Aguirre Verónica, Muñoz Rocío, Cáseres Germán, Medina Santiago, Dell'Oso Matías, Paniego Juan Manuel, Pi Puig Martín, Rodriguez Eguren, Sebastián. 2017. XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires). ISBN: 978-987-42-5143-5
11. Which Change Sets in Git Repositories Are Related?. Ramadani, J., Wagner, S. EEE 2016. International Conference on Software Quality, Reliability and Security (Viena, Austria.)
12. Synchronization and replication in the context of mobile applications. STAGE, A. (2005 Joint Advanced Student School Course 6: Next-Generation User-Centered).
13. III-LIDI: http://weblidi.info.unlp.edu.ar/wp/proyectos/investigacion/

Cambiando Percepciones: UX Design

Tatiana Hotimsky [2], Walter Molina [2], Carlos Salgado [1], Mario Peralta [1], Alberto Sánchez [1]
[1] Departamento de Informática de la Facultad de Ciencias Físico Matemáticas y Naturales
Universidad Nacional de San Luis
Ejército de los Andes 950, CP 5700, San Luis, Argentina
email: {csalgado, mperalta, alfanego} @unsl.edu.ar

[2] Departamento de Ciencias Aplicadas; Escuela de Gestión de Empresas y Economía;
Escuela de Ingeniería y Biotecnología
Universidad Nacional de Villa Mercedes
Las Heras 383, CP 5730, Villa Mercedes, San Luis, Argentina
email: {tatianahotimsky, waltermolina} @unvime.edu.ar

RESUMEN

Pensar en la experiencia que tendrá el usuario al interactuar con nuestra pieza de software, es comenzar a entender que en el proceso de interacción influyen distintos factores: individuales, sociales, culturales contextuales y aquellos propios del producto en cuestión. Esto será determinante en la percepción -positiva o negativa- que tendrá el usuario de un producto, servicio y/o proceso.

Es así, como se evidencia la necesidad de contar con alguna estrategia, técnica, modelo o método que permita organizar y/o sistematizar estas tareas a los desarrolladores de software y los profesionales de distintas disciplinas que trabajan en conjunto, de manera interdisciplinaria, en el desarrollo de interfaces de software.

Para el desarrollo del modelo propuesto se tuvo en cuenta:

1. La estructura de la Norma Internacional ISO 13407:1999 que describe cómo un proceso de Diseño Centrado en Usuario puede ser implementado para lograr sistemas usables; y su revisión ISO 9241-210:2010 confirmada en 2015.
2. Información recabada de la Industria del Software y Servicios Informáticos de la región.

El objetivo de nuestra propuesta es brindar a las empresas de desarrollo de software una herramienta, guía o buenas prácticas que les permita posicionarse en un nivel altamente competitivo en el mercado actual, mediante la producción de software de calidad basado en la experiencia de usuario.

PALABRAS CLAVE

Experiencia de Usuario, Usabilidad, Interacción, Diseño Centrado en el Usuario.

CONTEXTO

El presente trabajo se enmarca en el Proyecto de Investigación N°P-031516, denominado "Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software con Calidad" de la Facultad de Ciencias Físico Matemáticas y Naturales de la Universidad Nacional de San Luis. Este 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

Muchos de los críticos de la *experiencia de usuario* toman una frase atribuida a Henry Ford, hace más de 100 años, para poder fundamentar sus críticas:

> *"Si se hubiera preguntado a los clientes qué querían, hubieran respondido caballos más veloces."*

Si lo escuchamos o leemos al pasar podríamos creer que es una frase aplicable a la *experiencia de usuario* y supondríamos que ésta se trata de

los deseos de los usuarios y de cómo satisfacerlos. En ese caso, al momento de iniciar un desarrollo, nos preguntaríamos:

¿qué desean los usuarios?

Y así comenzaríamos una carrera por descubrir esa fórmula para generar un producto digital que tenga una conjunción que satisfaga los deseos de nuestros usuarios. Misión imposible.

Si siguiéramos ese camino, nos estaríamos olvidando de la esencia del *diseño centrado en el usuario*: entender al usuario en su contexto [1].

Entonces, la pregunta que debemos realizar no es *¿qué desea el usuario?* sino *¿cómo podemos mejorar la experiencia de los usuarios haciendo un uso eficiente de los recursos con los que contamos?*

Para entender mejor este interrogante es necesario remontarnos varios años atrás, y encontrarnos con el término *interacción persona-ordenador* (en adelante HCI, por sus siglas en inglés):

> *"Para poder cubrir todos los puntos de la definición, la HCI ha de abarcar diferentes áreas que incluyan aspectos del ser humano y del ordenador. Entre ellas destaca la informática, el diseño industrial, la psicología cognitiva y la sociología"* [2]

La HCI estudia [2]:

- el hardware y el software y como afectan a la interacción;
- los modelos mentales de los usuarios frente al modelo de la máquina;
- las tareas que desempeña el sistema y su adaptación a las necesidades del usuario;
- el diseño, que debe estar dirigido al usuario y no a la máquina (*user-centered design*);
- y el impacto organizacional, que deberá ser positivo.

Podemos ver que el término incorpora en la industria informática conceptos que se desprenden de la psicología, sociología y el diseño industrial; que tienen que ver con un aspecto subjetivo como pueden ser *los modelos mentales de los usuarios* o *las necesidades que tienen.*

Unos años después de que naciera este término surge, de la mano de la democratización de la información, en la década de los '90, el concepto *ingeniería de la usabilidad*; que propone una visión más estratégica y económica de generar una *buena* experiencia de usuario. Se centra en el ROI (Return On Investment o Retorno sobre la Inversión), la obtención de resultados y la relación costo-beneficio [3].

El concepto de *usabilidad* evoluciona con los años hasta convertirse en un método para valorar la *calidad* de un diseño interactivo. Y es definido de la siguiente manera en el estándar ISO/IEC 25010 [4]:

> *Capacidad del producto software para ser entendido, aprendido, usado y resultar atractivo para el usuario, cuando se usa bajo determinadas condiciones. Esta característica se subdivide a su vez en las siguientes subcaracterísticas:*
>
> - *Capacidad para reconocer su adecuación. Capacidad del producto que permite al usuario entender si el software es adecuado para sus necesidades.*
> - *Capacidad de aprendizaje. Capacidad del producto que permite al usuario aprender su aplicación.*
> - *Capacidad para ser usado. Capacidad del producto que permite al usuario operarlo y controlarlo con facilidad.*
> - *Protección contra errores de usuario. Capacidad del sistema para proteger a los usuarios de hacer errores.*
> - *Estética de la interfaz de usuario. Capacidad de la interfaz de usuario de agradar y satisfacer la interacción con el usuario.*
> - *Accesibilidad. Capacidad del producto que permite que sea utilizado por usuarios con determinadas características y discapacidades.*

Gavin Allanwood y Peter Beare plantean que *una vez terminado o, mejor aún, durante el proceso de diseño* de una aplicación de software*, las pruebas de usabilidad son una herramienta útil que puede emplearse para descubrir qué se debe mejorar"* [5]. Las pruebas de usabilidad tienen 5 componentes:

1. la capacidad de aprender,
2. la eficiencia,
3. la memorabilidad,

4. los errores,
5. la satisfacción.

Preguntarnos *¿qué debemos mejorar?,* no es lo mismo que preguntarnos *¿cómo podemos mejorar la experiencia de los usuarios haciendo un uso eficiente de los recursos con los que contamos?* El primer interrogante no incorpora la experiencia que tendrán los usuarios.

Existen reglas que definen la *experiencia de usuario* y la diferencian de la *usabilidad*. La *experiencia de usuario* según Marc Hassenzahl [6]*:*

1. Se enfoca en aspectos positivos de la relación usuario-producto.
2. Incorpora aspectos hedónicos.
3. Se enfoca en entender y administrar aspectos subjetivos del uso del producto.

Y así, el modelo hedónico/pragmático incorpora el concepto de *percepción de la interacción* que podrá ser:

1. Pragmática. Hace referencia a la capacidad de percepción del producto para respaldar el logro de los objetivos.
2. Hedónica. Se refiere a la capacidad percibida del producto para apoyar el logro de los objetivos principales, donde se encuentra el "por qué" de la acción que intenta completar el usuario.

Esta visión de la *experiencia de usuario* plantea un principio, que es el siguiente:

> "*No es antagónico distinguir entre una descripción completa de una experiencia y la abstracción de una meta-experiencia"*

De esta forma se logra diferenciar la *experiencia de usuario* de la *usabilidad,* pero a su vez se comprende que la UX forma parte del *diseño centrado en el usuario*.

Siendo parte del diseño centrado en el usuario, no podemos evitar recorrer las bases que plantea en este ámbito el World Wide Web Consortium -en adelante W3C-, a través de la Web Accessibility Initiative (WAI). En 1999 publicó, en el documento Web Content Accessibility Guidelines (WCAG), 14 principios con el objetivo de promover la accesibilidad haciendo foco no solamente en las capacidades de cada usuario, sino también en los dispositivos utilizados para consumir el contenido [7].

La revisión de 2008, WCAG 2.0,-que luego fue aprobada como estándar [8] en la norma ISO/IEC 40500:2012, establece 4 principios -con sus respectivas guías- para lograr la accesibilidad del contenido [9]:

1. Perceptibilidad,
2. Operabilidad,
3. Comprensibilidad,
4. Robustez.

Estos principios no han sido escritos para una tecnología en particular, lo que los hace generalizables para todos los desarrollos de interfaces de software. La guía WCAG2ICT describe cómo aplicarlos en otras plataformas [10].

En 2018, se publicó una nueva revisión, WCAG 2.1 que intenta resolver cuestiones propias de nuevas tecnologías disponibles [11].

Si analizamos en profundidad los 4 principios podemos notar como el Diseño Centrado en el Usuario está directamente vinculado con la responsabilidad del equipo de UX de entender al usuario y con conocer de forma certera su contexto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Desarrollar un producto digital sin preguntarnos quiénes lo usarán, para qué, desde qué dispositivos, etc., no hace más que alejarnos de la premisa que reina hoy en día, en esta era digital: todo por el usuario.

En el marco de esta línea de trabajo se llevó a cabo un Workshop, en el XIII Encuentro Latinoamericano de Diseño de la Universidad de Palermo, en el que se trabajó con una adaptación de la actividad propuesta en el libro Diseño de Experiencias de Usuario [5] denominada Ciclo de Desarrollo Centrado en el Usuario, que "introduce la idea de un proceso de diseño iterativo centrado en un grupo de usuarios y en un contexto específico" a través del diseño de una combinación de 3 platos de comida y en 4 etapas.

Este trabajo es una las validaciones que son necesarias para poder hablar de un modelo de calidad centrado en la experiencia del usuario.

Al finalizar la actividad se solicitó a los diferentes equipos de trabajo plantear, en conjunto, respuesta a las siguientes preguntas:

- ¿Conocer a los *potenciales usuarios* modificó la visión preliminar del menú?
- ¿Y conocer el *contexto*?
- ¿Qué hubiera pasado si se trataba de un producto digital?
- ¿Hubiera sido fácil introducir modificaciones con la aprobación de todo el equipo?
- ¿Hubiera sido fácil introducir cambios y/o modificaciones sobre una versión que ya está en etapa de prueba?
- ¿El proceso sería rentable económicamente hablando?

A menudo en la industria del software nos encontramos con productos digitales que son ideados y desarrollados por personas expertas en el tema, a las que resulta fácil entender el uso de un *menú hamburguesa* o que conviven en lo cotidiano con términos como *scroll, input* o *form.* Pero ¿qué pasa si hablamos de usuarios que no son expertos en el tema? ¿Y si hablamos de los diferentes grupos etarios? ¿De comunidades vulnerables o que no poseen fácil acceso a la tecnología? ¿Cómo se sienten al encontrarse con un producto digital que no fue ideado o desarrollado para su situación o su contexto?

Ahora preguntémonos: ¿cómo podrían los equipos de desarrollo sacar al mercado un producto digital pensando en el usuario si generalmente trabajan contra reloj? Si, a menudo conocen al usuario en el transcurso del desarrollo, ¿podrían cubrir las modificaciones que van siendo solicitadas? A las respuestas sumémosle que los equipos están conformados por personas con diferente visión, y que pueden pensar diversas soluciones según su experiencia ¿Cómo se podría lograr un consenso?

Y si tenemos en cuenta que, a menudo, los equipos de desarrollo deben producir software multiplataforma, impactando directamente sobre la UX por los riesgos propios de las tecnologías que se usan [12], ¿podríamos llegar a un producto digital con el que todos, equipo y usuarios, se sintieran satisfechos e identificados?

Reflexionando sobre las respuestas que fueron surgiendo, se obtuvo la siguiente conclusión: realizar los pasos de la actividad propuesta, pero de atrás para adelante hubiera facilitado el desarrollo y puesta en marcha del menú, optimizando el uso de los recursos (humanos, económicos, materiales, etc.) y logrando una *experiencia de usuario* memorable.

Atendiendo a esta reflexión se pueden proponer los siguientes pasos para el proceso de diseño de una solución centrada en el usuario.

1. Analizar el problema y tomarnos el tiempo para preguntarnos en qué contexto será utilizado el producto y conocer a quiénes lo utilizarán. Plantear interrogantes a tener en cuenta que nos permiten proponer ideas.
2. Conocer el contexto en profundidad, preguntarnos cuál de las propuestas que tenemos pensadas e ideadas será la más adecuada y óptima.
3. Comprender al usuario y sus características. Ponernos en sus zapatos. Hacer testeos con la opción seleccionada para desarrollo, y así verificar si la decisión fue la adecuada.
4. Conociendo el contexto y a los *potenciales usuarios* y habiendo realizado una prueba preliminar, esta será una instancia para pulir las propuestas que fueron surgiendo a lo largo del proceso. Teniendo en claro que lograremos una excelente experiencia de usuario, mejorando así la calidad de nuestro producto digital.

Para ello, se está trabajando en la definición de modelos, métodos, guías, estrategias que permitan evaluar la calidad de las interfaces con foco en las experiencias de usuario.

3. RESULTADOS OBTENIDOS

Luego de conocer las consideraciones obtenidas en el workshop y de analizar la perspectiva de diferentes autores, podemos comprender que conocer el contexto y a los usuarios que harán uso de nuestro producto digital nos dará la posibilidad de delimitar nuestro campo de diseño y desarrollo, optimizando esfuerzos y recursos, para obtener un producto de calidad asegurando mejorar la relación usuario-producto.

Al finalizar el desarrollo de un producto digital, podríamos plantearnos algunos interrogantes:

- ¿Le aporta valor al usuario?
- ¿Es perceptible, operable y compresible?
- ¿El usuario tiene una experiencia positiva?
- ¿Logramos un producto final robusto?

Si la respuesta a todas esas preguntas es ¡SÍ!, tendremos la certeza de estar frente a una experiencia de usuario memorable. Y eso es, en otras palabras, empezar a entender que, si hablamos de calidad en el área del software, hablamos de experiencia de usuario.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación, se logró un trabajo conjunto entre la Universidad Nacional de San Luis (UNSL) y la Universidad Nacional de Villa Mercedes (UNViMe).

Por parte de la UNSL, el equipo está compuesto por 3 docentes investigadores categorizados a nivel nacional, 2 tesistas de posgrado de la Maestría en Ingeniería de Software y 1 becario de la Ingeniería en Informática. En este marco se están llevando a cabo 2 tesis de posgrado y algunos trabajos finales de carrera de la Ingeniería en Informática y tesinas de la Licenciatura en Ciencias de la Computación.

Desde la UNViMe, la formación de recursos humanos incluye a 2 de los autores de este documento, docentes de la universidad, quienes son maestrandos de la Maestría en Calidad de Software de la UNSL.

5. BIBLIOGRAFÍA

[1] International Organization for Standardization, ISO 9241-210:2010 Ergonomics of human-system interaction -- Part 210: Human-centred design for interactive systems, 2010.

[2] M.-C. Marcos, «HCI (human computer interaction): concepto y desarrollo,» *El Profesional de la Información,* vol. 10, nº 6, 2001.

[3] J. Nielsen, Usability Engineering, Mountain View, California, 1993.

[4] International Organization for Standardization & International Electrotechnical Commission, ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models, 2011.

[5] G. Allanwood y P. Beare, Diseño de Experiencias de Usuario, Primera ed., Badalona (Barcelona): Parramón Arts & Design, 2015, pp. 12-36.

[6] M. Hassenzahl, *The Hedonic/Pragmatic Model of User Experience,* Landau, 207.

[7] World Wide Web Consortium, «Web Content Accessibility Guidelines 1.0,» 5 Mayo 1999. [En línea]. Available: http://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505. [Último acceso: 9 3 2019].

[8] International Organization for Standardization & International Electrotechnical Commission, ISO/IEC 40500:2012 Information technology -- W3C Web Content Accessibility Guidelines (WCAG) 2.0, 2012.

[9] World Wide Web Consortium, «Web Content Accessibility Guidelines (WCAG) 2.0,» 11 Diciembre 2008. [En línea]. Available: http://www.w3.org/TR/2008/REC-WCAG20-20081211/. [Último acceso: 9 Marzo 2019].

[10] World Wide Web Consortium, «Guidance on Applying WCAG 2.0 to Non-Web Information and Communications Technologies (WCAG2ICT),» 5 Septiembre 2013. [En línea]. Available: http://www.w3.org/TR/2013/NOTE-wcag2ict-20130905/. [Último acceso: 9 Marzo 2019].

[11] World Wide Web Consortium, «Web Content Accessibility Guidelines (WCAG) 2.1,» 5 Junio 2018. [En línea]. Available: https://www.w3.org/TR/2018/REC-WCAG21-20180605/. [Último acceso: 9 Marzo 2019].

[12] A. Biørn-Hansen, T.-M. Grønli, G. Ghinea y S. Alouneh, «An Empirical Study of Cross-Platform Mobile,» *Wireless Communications and Mobile Computing,* vol. 2019, nº 5743892, 3 Enero 2019.

[13] T. L. P. P. M.G. Helander, Handbook of Human-Computer Interaction, 1997.

[14] International Organization for Standardization, ISO 13407:1999, 1999.

[15] International Organization for Standardization, ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability, 1998.

[16] International Organization for Standardization, ISO 9241-11:2018, 2018.

Coreografías de servicios: una adaptación para dispositivos ubicuos

Germán Montejano [1 2]; Oscar Testa[2]; Rubén Pizarro[2]; Darío Segovia[2]

[1]Facultad de Ciencias Físico-Matematicas 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
[2] Facultad 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

Resumen

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. 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 construcción 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 dispositivos[1] 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 distribuídos. Podría pensarse que son aplicables a ambientes ubicuos. Sin embargo, 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 dipositivos 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 quepermitan solucionar en todo o en parte estos desafíos planteados, haciendo

[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.

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: cámas 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 se encuentra en su fase finalización, donde nos encontramos analizando los resultados obtenidos de las 6 pruebas de concepto realizadas, para realizar los ajustes finales tanto en el framework de ejecucución (tanto en arduino como en PHP) así como los agregados a la especificación WS-CDL para dar soporte a las características específicas de los dispositivos utilizados. 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 otro 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 Geráis 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 tesinas 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. Hagras, M. Weber, and W. Minker, eds.), pp. 195–223, Springer US, 2009.

Datos abiertos en Educación, evaluación de los alcances de las iniciativas existentes. Caso de estudio Universidad Nacional de Catamarca

Herrera, Claudia M., Chayle, Carolina I., Barrera, María A., Pauletto, Ana C., Leguizamón Almendra; Juan C. y Morales Apaza, Oswaldo W. N.

Departamento de Informática/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
Maximio Victoria Nº 55 - C.P: 4700 - San Fernando del Valle de Catamarca
Telefono: 03833- 435112 – int 168
cchayle@gmail.com, herrera.claudia.mabel@gmail.com

Resumen

En los últimos años ha comenzado una verdadera revolución en materia de acceso a la información pública por parte de los ciudadanos. A esta revolución se le ha denominado “Open Data” o “Datos Abiertos” (DA). Este nuevo modelo Este nuevo modelo se basa en el acceso y uso de la información pública por parte de terceros para entregar nuevos servicios a los ciudadanos. El concepto de gobierno abierto introduce una nueva forma de relación entre la sociedad, el gobierno y la administración, en el que las Tecnologías de la Información (TI) cobran un papel central. Si se trasladan los principios de gobierno abierto al ámbito universitario, surge la idea de “Universidad Abierta”. La apertura de datos en los diferentes dominios que articulan la universidad (docencia, investigación, biblioteca, archivo, etc.) aporta notables beneficios tanto a la comunidad universitaria y a la universidad como institución, como a la propia sociedad donde se inserta.

La presente investigación tiene por objeto evaluar algunos de los alcances tempranos de las iniciativas existentes en el tema de datos abiertos y educación, en el ámbito de la Universidad Nacional de Catamarca (UNCA), produciendo una contribución efectiva al proyecto de investigación denominado "Las Tic al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública" que se ejecuta en el Departamento de Informática.

PALABRAS CLAVES: Dato Abierto, Universidad Abierta, TIC, UNCA.

Contexto

Se presentan los avances vinculados a los alcances tempranos de las iniciativas existentes en el tema de datos abiertos y educación, en el ámbito de la Universidad Nacional de Catamarca (UNCA), enmarcados en el proyecto “Datos abiertos en Educación, evaluación de los alcances de las iniciativas existentes. Caso de estudio Universidad Nacional de Catamarca”, perteneciente a los Proyectos de Investigación y Desarrollo de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA, produciendo una contribución efectiva al proyecto de investigación denominado "Las Tic al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública" que se ejecuta en el Departamento de Informática.

1. Introducción

En una nueva etapa de la evolución del gobierno electrónico se comenzó a utilizar la Web como un canal paralelo para el acceso a los servicios por parte del ciudadano, existiendo una presencia masiva de

organizaciones gubernamentales en la Web, y ofreciéndose servicios como por ejemplo: llenado y envío de formularios electrónicos u ofreciéndose transacciones completas y seguras tales como, pago de impuestos, gestión on line de trámites, etc. En los últimos años ha comenzado una verdadera revolución en materia de acceso a la información pública por parte de los ciudadanos. A esta revolución se le ha denominado "Open Data" o "Datos Abiertos" (DA). Este nuevo modelo consiste en poner a disposición de la sociedad los datos de interés común de la ciudadanía y en el uso de la información pública por parte de terceros para entregar nuevos servicios a los ciudadanos.

Los datos abiertos son aquellos a los que cualquier persona puede acceder, usar y compartir libremente. De esta manera, la publicación de los datos en formatos abiertos es la mejor forma de generar confianza en las instituciones, gracias a la transparencia que esto aporta al sistema, como a la exposición pública del trabajo en el que se invierten los recursos públicos (principalmente los recursos económicos). Este arduo trabajo, de generar, publicar, ordenar datos en abierto, se está convirtiendo en una realidad activa. El avance de la informática y las TIC ha permitido la detonación de los datos gracias a la reducción de costos y de tiempos en su publicación y sociabilización, por eso la apertura de datos está alcanzando tal atención e importancia en estos momentos. El concepto de gobierno abierto introduce una nueva forma de relación entre la sociedad, el gobierno y la administración, en el que las Tecnologías de la Información (TI) cobran un papel central. Si para los gobiernos, la publicación de datos abiertos es un imperativo, para las universidades –en especial las públicas– su adopción emerge no sólo como un deber, sino como una oportunidad de extender su quehacer con la comunidad y establecer vínculos con otras instituciones de educación superior, beneficiándose también, de esta manera, de la reutilización de los datos de investigación científica. Las universidades, como instituciones públicas, tienen una responsabilidad social elevada. De hecho, debido a que una de las misiones de la Universidad es contribuir, de manera general, al bienestar de la sociedad, es preciso considerar los conceptos de transparencia y datos abiertos.

Por tanto, las universidades deben desempeñar diferentes acciones a favor de la transparencia para consolidarse como organismos abiertos de cara a la ciudadanía.

Si se trasladan los principios de gobierno abierto al ámbito universitario, surge la idea de "Universidad Abierta". Una universidad abierta es aquella que favorece el acceso abierto a la información, escucha a la comunidad universitaria, tomando decisiones según sus necesidades y prioridades, y genera espacios de colaboración para desarrollar servicios con todos los actores de la comunidad universitaria.

Si se trasladan los principios de gobierno abierto al ámbito universitario, surge la idea de "Universidad Abierta". Finalmente, la posibilidad de hacer uso intensivo y planificado de las TI juega un papel fundamental en el concepto de Universidad Abierta, ya que son precisamente éstas las que ayudan a hacer realidad los tres pilares fundamentales de la apertura de datos: disponibilidad y acceso; reutilización y redistribución; y participación universal.

Además, la apertura de datos en los diferentes dominios que articulan la universidad (docencia, investigación, biblioteca, archivo, etc.) aporta notables beneficios tanto a la comunidad universitaria y a la universidad como institución, como a la propia sociedad donde se inserta.

Este carácter abierto ayudará a ampliar las fronteras del conocimiento y mejorará las bases de transparencia, colaboración y participación de la sociedad del futuro.

En Argentina estos principios están siendo ampliamente respaldados a través de la legislación promulgada. En el año 2016, mediante el Decreto 117/2016 (12-01-2016), el gobierno nacional impulsó el "Plan de

Apertura de Datos" el cual involucra a todo el Poder Ejecutivo y designa como responsable de su implementación al Ministerio de Modernización. Esta legislación formalizó la decisión del Poder Ejecutivo de abrir los datos públicos y consolidar el Portal de Datos de la Nación. Este Decreto, y su cumplimiento, representa un gran paso para la profundización de uno de los ejes necesarios para la consolidación de un Gobierno Abierto en la Argentina.

Por otra parte, con fecha 14-09-2016 se sanciona la Ley Nº 27.275 de Derecho de Acceso a la Información Pública que tienen por objeto garantizar el efectivo ejercicio del derecho de acceso a la información pública, promover la participación ciudadana y la transparencia de la gestión pública; y establece que la información en poder del Estado debe ser accesible para todas las personas y estar disponible en formatos electrónicos abiertos para facilitar su circulación y redistribución.

Actualmente, en distintas universidades de América Latina se plasman iniciativas de datos abiertos, permitiendo a sus comunidades educativas que tengan acceso a datos que van desde información institucional hasta repositorios de investigación; como así también existen universidades que han ejecutado los proyectos de gobierno en línea y abierto en su totalidad. Iniciativas similares se detectan en universidades de la República Argentina tales son los casos de la Universidad Nacional de Cuyo y el programa de Universidad Abierta de la Universidad de Buenos Aires.

La presente investigación tiene por objeto evaluar algunos de los alcances tempranos de las iniciativas existentes en el tema de datos abiertos y educación, en el ámbito de la Universidad Nacional de Catamarca (UNCA), produciendo una contribución efectiva al proyecto de investigación denominado "Las Tic al Servicio del Dato Abierto: Situación Actual, Conceptualización e Iniciativas de Apertura de Información Pública" que se ejecuta en el Departamento de Informática.

2. Líneas de Investigación, Desarrollo e Innovación

El presente proyecto es un investigación aplicada en la cual se realizarán estudios de carácter descriptivo y comparativo, se analizarán los sitios Web de la Universidad Nacional de Catamarca, con el fin de determinar los alcances de las iniciativas de datos abiertos en el ámbito de la Universidad Nacional de Catamarca. Los parámetros seleccionados para confeccionar la evaluación serán, por un lado realizar un relevamiento de la utilización de ejes asociados a Universidad Abierta y, por el otro, analizar los compromisos en materia de apertura de datos (open data), ambos en el ámbito de la Universidad Nacional de Catamarca.

Con respecto a la obtención de la información se prevé recolección de muestras mediante la observación directa a los sitios que pertenezcan a la UNCA. Las metodologías, técnicas y herramientas serán exhaustivamente analizadas para elegir la que mejor se adapte a la evaluación de las iniciativas de dato abierto existentes.

3. Resultados y Objetivos

A partir de la sanción de la Ley Nacional Nº 27.275 de Derecho de Acceso a la Información Pública y del Decreto 117/2016, mediante el cual el gobierno nacional impulsó el "Plan de Apertura de Datos" y de la evaluación de los alcances de las iniciativas de datos abiertos en el ámbito de la UNCA, se tratará de responder a la pregunta ¿Qué se ha logrado hasta ahora con la apertura de datos?

Con la legislación nacional vigente en materia de acceso a la información, se hace necesario jerarquizar como bien público la información producida por la universidad y garantizar el acceso a toda documentación generada en su ámbito; por lo que se tratará de brindar recomendaciones para la apertura de datos en la UNCA, que apunten a transparentar aún más la gestión universitaria y facilitar el

acceso a la información.

Los objetivos del proyecto son:

- Evaluar los alcances de las iniciativas de datos abiertos en el ámbito de la Universidad Nacional de Catamarca (UNCA).
- Estudiar los casos de universidades abiertas existentes en América Latina.
- Observar las aperturas de datos de educación, realizadas en el ámbito de las Universidades Argentinas.
- Analizar las aperturas de datos de educación, en el ámbito de la UNCA.
- Caracterizar los datos abiertos de educación de la UNCA.
- Recomendar acciones a seguir en el diseño de proyectos de universidad abierta en el ámbito de la UNCA.
- Acercar a la universidad al cumplimiento de la Ley Nacional Nº 27.275 de Derecho de Acceso a la Información Pública y del Decreto 117/2016 (12-01-2016), mediante el cual el gobierno nacional impulsó el "Plan de Apertura de Datos".

4. Formación de Recursos Humanos

El proyecto está integrado por 1 (un) profesor titular, 2 (dos) profesores adjuntos, 2 (dos) docentes auxiliares, y 1 (un) alumno. El mismo tiene como objetivo la formación de recursos humanos a través de:

- Participación en eventos científicos regionales, nacionales e internacionales de la especialidad, como congresos, simposios, seminarios y cursos, por ello el programa de capacitación y formación de recursos humanos, contempla las siguientes actividades:
- Participación en cursos de actualización y posgrado en el área de estudio.
- Participación de los integrantes del proyecto en la dirección, asesoramiento y evaluación de tesinas de grado de la carrera de Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCA.
- Desarrollo de tesis de posgrados por parte de los docentes integrantes.

En cuanto a la formación específica del equipo de docentes investigadores, se destaca que los mismos se encuentran abocados a actividades de capacitación y estudios de posgrado, para la elaboración de 2 (dos) tesis correspondiente a la carrera Maestría en Ingeniería del Software de la Universidad Nacional de San Luis.

5. Referencias

Aparicio J. M., A. Fuster, I. Garrigós, F. Maciá, J. N. Mazón, L. Vaquer, J. J. Zubcoff.

Ecosistema de Datos Abiertos de la Universidad de Alicante - Universidad de Alicante.

Concha G. y A. Naser. Datos abiertos: Un nuevo desafío para los gobiernos de la región. Serie Gestión Pública. Instituto Latinoamericano y del Caribe de Planificación Económica y Social (ILPES). Santiago de Chile. Marzo de 2012.

Decreto 117/2016. Plan de Apertura de Datos. Argentina. 2016.

Ley Nº 27.275. Derecho de Acceso a la Información Pública. Argentina. 2016.

Ministerio de Modernización de la Presidencia de la Nación. Kit de Datos Abiertos.

Morales Vargas, A. CAPÍTULO 1 "Datos abiertos y visualización de información en sitios web de universidades chilenas: una asignatura pendiente". Visualización de la Información en el Campo del Periodismo y la Comunicación SociaL. Universidad de Chile.

Open Data Census. Índice de Datos Abiertos Ciudades de Argentina.

Palacios Rodriguez, S. "Open Data: Antecedentes en el mundo". OPEN DATA - Miradas y Perspectivas de los Datos Abiertos. Páginas: 24 a 27. Universidad Nacional de La Rioja. Septiembre 2015. 1º Edición.

Diseño de una metodología para requerimientos referenciales en dominios complejos

Moyano Ezequiel UNTDF, Urciuolo Adriana UNTDF, Czelada Alejandro UNTDF, Izarra Emilio UNTDF, Rigoni Brian UNTDF
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-432403
jmoyano@untdf.edu.ar, aurciuolo@untdf.edu.ar, eaczelada@untdf.edu.ar, eizarra@untdf.edu.ar, brian.rigoni1@gmail.com

Resumen

Los sistemas que modelan el comportamiento de la naturaleza se caracterizan por su complejidad, la mayoría desarrollados por expertos del dominio, sin considerar el uso de métodos y herramientas de Ingeniería de software; así el conocimiento adquirido queda solo en ese ámbito. Para el desarrollo de nuevos sistemas el conocimiento debe obtenerse nuevamente. El reuso del conocimiento del dominio resulta esencial.

Para lograr el reuso del conocimiento es necesario contar con métodos de especificación de requerimientos a nivel de dominio (Requerimientos Referenciales), que puedan ser utilizados en futuros sistemas del dominio, reduciendo tiempos y esfuerzos.

La ingeniería de dominio tiene entre sus objetivos el reuso de recursos centrales (coreassets) de un dominio específico, abarca procesos vinculados al modelado de dominio siendo el Análisis de Dominio una de sus principales etapas (base para el desarrollo de software reusable).

A pesar de la importancia de la etapa de requerimientos en el desarrollo de software, existen pocos antecedentes sobre métodos para la formulación de Requerimientos Referenciales como paso previo al Análisis de Dominio.

El Proyecto propone definir una metodología para la especificación de Requerimientos Referenciales en dominios complejos, para demostrar la importancia del reuso del conocimiento en un sistema concreto como caso de estudio (modelado de Cuencas)

Palabras clave: Requerimientos Referenciales, Ingeniería de Requerimientos, Análisis de Dominio, Reuso.

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la Univdersidad Nacional de Tierra del Fuego, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática de Sistemas de Información Ambiental – Hidroinformática desde hace más de 10 años.

El proyecto actual tiene como caso de estudio Aplicaciones de Modelado de Cuencas durante el período (Abr/2019-Mar/2021).

Introducción

El estudio de los sistemas que modelan la naturaleza adquiere cada vez mayor importancia, dada la función preponderante que los mismos cumplen en la toma de decisiones para el manejo y aprovechamiento sustentable de los recursos naturales y el medio ambiente. Los modelos ambientales son abstracciones de la naturaleza y de su comportamiento, que permiten la toma de decisión a partir de predicciones.

La necesidad de contar con información actual y exacta es alta, pero además de esto, la complejidad de estos sistemas está dada por la necesidad de manejar grandes volúmenes de datos, objetos espacio / temporales, factores ambientales, políticas de gestión, normativas

legales y técnicas. El reuso de conocimiento del dominio en tales situaciones resulta esencial, dado el esfuerzo que conlleva la adquisición de la información en las distintas organizaciones vinculadas al manejo ambiental.

Estos sistemas en gran parte fueron desarrollados por expertos del dominio en forma aislada, sin considerar el uso de métodos y herramientas de Ingeniería de software; con lo cual el conocimiento adquirido queda solo en el ámbito en el cual se obtuvo (organismo o institución). Para el desarrollo de nuevos sistemas (en el mismo dominio), el conocimiento debe obtenerse nuevamente, lo cual involucra un gran esfuerzo. De allí la importancia del reuso de conocimiento en dominios complejos.

Los sistemas de información ambiental, como los de modelado de cuencas, se caracterizan por su complejidad. Rozan una gran cantidad de subdominios diversos de conocimiento, generalmente heterogéneos, tanto técnica como semánticamente.

Estos sistemas se relacionan con el manejo de los datos correspondientes a los distintos componentes interactuantes del ambiente: el suelo, el agua, el aire y las especies existentes.

La mayoría de los fenómenos ambientales se debe modelar considerando cinco dimensiones: su localización en el espacio (latitud, longitud y altitud), el tiempo en que se obtiene y los fenómenos particulares que son analizados.

La meta de un control del medio ambiente y de un sistema de información es colectar así datos en este mundo cinco - dimensional, almacenarlos, analizarlos, y promover su uso de manera más eficiente. Esta meta requiere la *integración del espacio y del tiempo*, un campo complejo, donde todavía se necesita mucha investigación.

Normalmente el comportamiento de los sistemas ambientales se predice a través de modelos de simulación.

La ingeniería de dominio tiene como objetivo lograr la reutilización del conocimiento y de las especificaciones de un dominio específico, de manera de poder ser utilizadas en diversas aplicaciones del mismo ámbito.

Las técnicas de análisis de dominio están asociadas a la reutilización, su principal característica es capturar información relacionada con el dominio, y determinar qué y cómo este conocimiento se reutilizará en el desarrollo de futuras aplicaciones. El análisis del dominio produce un *modelo del dominio,* intentando encontrar los elementos más comunes de varios modelos de sistemas en el dominio.

No obstante existen escasas técnicas de especificación de requerimientos a nivel de dominio (Requerimientos Referenciales) que permitan, para sistemas de tal complejidad, obtener un modelo de dominio completo y consistente.

Se puede hacer un paralelismo entre la ingeniería de software tradicional y la ingeniería de dominio, así como la primer etapa en la ingeniera de software tradicional se llevan adelante los procesos (elicitación-especificación-validación) de la ingeniería de requerimientos, donde existen un gran desarrollo (previo a la etapa de análisis); en la ingeniería de domino, como paso previo al análisis de dominio, se deben obtener requerimientos referenciales.

El análisis de dominio se constituye en el punto de partida para obtener software reusable, sobre todo para sistemas de gran complejidad, facilitando y permitiendo el reuso de los productos obtenidos para las actividades futuras del desarrollo de software. La obtención de requerimientos referenciales o de

alto nivel resulta esencial para un análisis de dominio pertinente.

Para lograr el reuso del conocimiento es necesario contar con métodos de especificación de requerimientos a nivel de dominio (*requerimientos referenciales*), con el propósito de que la información (identificada, capturada y organizada) sea reutilizable en la creación de nuevas aplicaciones de software en ese dominio, en otros subdominios similares, reduciendo tiempos y esfuerzos de considerable magnitud.

La especificación de requerimientos, sobre todo en sistemas complejos, representa la comprensión de las necesidades de diferentes interesados del sistema (stakeholders) que se encuentran relacionadas con el mismo dominio. Muchas veces, los requerimientos son acordados por los interesados de tal forma que la semántica y significado de cada término utilizado es bien entendido; sin embargo cuando existen diferentes puntos de vistas del mismo concepto dentro de un dominio, ambigüedades e inconsistencias pueden surgir siendo perjudiciales para la Especificación de Requerimientos de Software.

A pesar de la importancia de la etapa de requerimientos en el desarrollo de software, existen pocos antecedentes sobre métodos para la formulación de requerimientos referenciales como punto de partida para el Análisis de Dominio, especialmente en dominio complejos donde adquiere más trascendencia para el reuso.

Si además se tiene en cuenta que los modelos ambientales (predicción de catástrofes, efectos del cambio climácico, sustentabilidad de los recursos hídricos, entre otros) se utilizan en distintas organizaciones de todo el mundo, obtener requerimientos referenciales posibilitará importantes avances para el desarrollo de modelos de dominio completos y consistentes, máxime su utilidad regional y global.

El principal problema es que no existen actualmente metodologías adecuadas para el análisis de requerimientos referenciales en estos dominios, que permitan obtenerlos y especificarlos de manera de desarrollar modelos de dominio y utilizarlos en diferentes aplicaciones. Resulta fundamental abrir nuevas investigaciones en su desarrollo.

La gran diversidad y heterogeneidad de los métodos, conceptos, aplicaciones y funcionalidades de los modelos existentes en el mundo, la posibilidad de contar con una sistematización del conocimiento del dominio es fundamental.

Línea de Investigación y Desarrollo

Como parte de la línea de investigación que el equipo de investigadores lleva adelante en temáticas vinculadas a los Sistemas de Información Ambiental – Hidroinformática, se plantean los siguientes objetivos a alcanzar en el presente proyecto:

Objetivo General

Promover el reuso de conocimiento en dominios complejos a través del desarrollo de una metodología de análisis y especificación de Requerimientos Referenciales, que permita reducir tiempo y esfuerzo en la construcción de aplicaciones de un dominio de características complejas.

Objetivos Específicos

1- Sintetizar el estado del arte respecto del análisis de requerimientos referenciales en dominios complejos.

2- Definir una metodología para la especificación de requerimientos referenciales en dominios complejos.

3- Aplicar la metodología elaborada previamente, utilizando como caso de estudio (Modelado de Cuencas).

4- Validar que los Requerimientos Referenciales obtenidos permitan el desarrollo de un modelo de dominio completo y consistente en modelos de simulación de cuencas, utilizando dos aplicaciones.

Hipótesis

Como hipótesis general de la investigación se plantea que desarrollando una metodología para la especificación de requerimientos referenciales, como paso previo al análisis de dominio, se podrá obtener un modelo de dominio completo y consistente que facilitará el desarrollo de distintas aplicaciones en el dominio específico.

Como sub-hipótesis se plantea tomar como caso de estudio el subdominio de aplicaciones de Modelado de Cuencas (como dominio complejo), que permita validar la metodología diseñada.

Resultados Esperados

Los resultados del proyecto aportarán la dinámica y disponibilidad de obtener requerimientos referenciales en dominios complejos, promoviendo el reuso del conocimiento obtenido en futuros desarrollos dentro del dominio, en particular en el modelado de cuencas (caso de estudio) donde la información es escasa.

Para lo cual se espera lograr:

- Redefinir los procesos de la ingeniera de requerimientos tradicional (adquisición, especificación, validación y documentación) a efectos de readecuarlos al modelado de un dominio complejo.
- Definir una nomenclatura y/o técnica (o ampliar alguna existente) para clasificar y especificar los requerimientos referenciales.
- Incorporar extensiones (a la técnica definida) necesarias que permita identificar y especificar los aspectos comunes, o similitudes (commonalities) y puntos de variación (variabilities) que contienen los diferentes sistemas que se pueden desarrollar en un dominio especifico
- Diseñar una estrategia que permita administrar la volatilidad de los requerimientos con el paso del tiempo.

Se espera que el impacto del proyecto sea la implementación de una metodología que pueda aplicarse a cualquier dominio cuya principal característica sea la complejidad.

Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF, licenciados en informática y alumnos de la Carrera.

El Proyecto incluye docentes en etapa de formación de postgrado, el Director del proyecto está realizando actualmente su tesis de magister en Ingeniería de Software en la Universidad Nacional de La Plata, denominada: “Reuso de Requerimientos Referenciales en Dominios Complejos. Caso de Estudio: Aplicaciones de Modelado de Cuencas.” Dirigida por la Mgs. Adriana Urciuolo y co-dirigida por el Dr. Gustavo Rossi de la UNLP.

Participaran Asistentes principales que realizarán sus primeras experiencias en actividades de investigación y que podrán desarrollar sus estudios de postgrado en temáticas afines al proyecto (Alejandro Czelada y Emilio Izarra).

Colaborará y prestará apoyo un alumno (Brian Rigoni) en formación de grado, con el objetivo formarlo en los aspecto generales de la investigación y que le sirva para el desarrollo de su futura tesis de grado.

Por otra parte, a través del proyecto se espera consolidar un equipo de trabajo de la UNTDF en la temática, sumando nuevos integrantes al

grupo principal (el cual ya ha ejecutado numerosos proyectos vinculados al problema).

Este proyecto permite además consolidar la cooperación con otros organismos. A su vez los docentes podrán nutrirse de la experiencia de quienes manejan diariamente los problemas y conflictos vinculados al manejo de dominios complejos como el caso de modelado de cuencas.

Referencias

[1] *Lenguajes Específicos de Dominio (DSL) para la Modelación de Ecosistemas Naturales.* Urciuolo Adriana, Gel Matías, Iturraspe Rodolfo, Moyano, Ezequiel, Villarreal Martín, Instituto de Desarrollo Económico e Innovación, UNTDF.

[2] *Aplicando herramientas MDE en la definición de un lenguaje especifico de dominio para la gestión de modelos*. Gabriela Pérez, Jerónimo Irazábal, Claudia Pons y Roxana GiandiniLIFIA, Facultad de Informática, Universidad Nacional de La Plata.

[3]*Cofessions of a Used Program Salesman: Intitutionalizing Software Reuse* Will Tracz..Addison Wesley, 1995.

[4]*Nuevo enfoque para la enseñanza del paradigma MDD: Ingeniería de Requerimientos basado en Modelos*Leopoldo Nahuel, Cecilia Ariste y Roxana Giandinil UNLP 2016.

[5]*Model-Driven Domain Analysis and Software Development*JanisOsis, Erika Asnina Riga Technical University, Latvia. 2010.

[6] *A Feature Modeling Approach for Domain - Specific Requirement Elicitation.* Olga De Troyer and Erik Janssens Dpt. Computer Vrije Universiteit Brussel.

[07] Luis Martíne*, Un Modelo De Mediación para el Desarrollo de Software Basado en Componentes COTS, Universidad De Málaga, Julio 2003.*

[08] Vanessa Hamar *Aspectos Metodológicos Desarrollo y Reutilización de Componentes de Software, Mérida – 2003.*

[09] Emeline, Regis; Gustavo, Tondello; *Análise de Domínio,(Universidad e Federal de Santa Catarina, 2000)*

[10] Urciuolo Adriana, Iturraspe Rodolfo, Moyano Ezequiel. *Perfil UML 2.0 para Aplicaciones de Monitoreo Ambiental.* VI JIISIC'07, Lima, Perú. Facultad de Cs e Ing., Pontificia Universidad Católica del Perú 2007, ISBN 978-9972-2885-1-7. pp. 393-401.

[11] Urciuolo A., R. Iturraspe, *"Conceptual Patterns for Water Resources Information Systems*", En: *Journal of Computer Science and Technology* Vol. 3 - No. 1 - April 2003 - ISSN: 1666-6038, pp 20-26.

[12] Moyano Ezequiel *Tesis de Grado: Técnicas de Análisis de Dominio para el Desarrollo de Componentes en Sistemas Complejos, Caso de Estudio: Sistemas de Información Ambiental, dirigida por* Urciuolo A. 2007.

[13]*Entorno de ingeniería de requisitos aplicado para producir software en una universidad*,L. Torres Pérez, M. D. Delgado Dapena, D. Rodríguez Nápoles, D. Gómez Suárez, W.De la Torre Parejo y Y. Alonso Abreu, 2014.

[14] *An Effective Requirement Engineering Process Model for Software Development and Requirements Management*. Dhirendra, U. S.:. In: International Conference on Advances in Recent Technologies in Communication and Computing (2010)

[15] *Comparing the performance of quantum-inspired evolutionary algorithms for the solution of software requirements selection problem*. Charan, K. S.Elsevier (2016)

[16] *Soft competency requirements in requirements engineering, software design, implementation and testing.* Holtkamp, P. Elsevier (2014)

[17] *User Requirements modeling and analysis of software-intensive systems*. Dos Santos, M, Vrancken, J Elsevier (2010)

DISEÑO Y EVALUACIÓN DE EXPERIENCIA DE USUARIO (UX) PARA MULTI-DISPOSITIVOS

Adriana MARTIN, Gabriela GAETAN, Viviana SALDAÑO, Claudia CARDOZO, Alejandra CARRIZO, Silvia VILLAGRA

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// acarrizo// svillagra }@uaco.unpa.edu.ar;
claudia_yoryi@yahoo.com.a

RESUMEN

Hace algunos años atrás, el diseño de productos estaba guiado por sólo dos tipos de dispositivos y sus respectivas plataformas: las computadoras y los celulares. Hoy, las personas eligen y utilizan una gran variedad de dispositivos para acceder e interactuar con la Web y mantener conexión con el mundo, tales como, PCs, teléfonos inteligentes, *tablets*, televisores, relojes y pulseras multi-propósito, etc. Estos dispositivos se pueden relacionar entre sí, compartiendo sus capacidades tecnológicas y la posibilidad de alternar entre ellos, en la continuidad de tareas y la concreción de los objetivos de sus usuarios. Haciendo uso de estas capacidades, las personas mantienen la interacción desde diversos dispositivos con las aplicaciones de su preferencia. Sin embargo, no es tan frecuente, como sería deseable, que las aplicaciones se conciban para una experiencia multi-dispositivo, que considere además a los usuarios destinatarios de las mismas.

El Proyecto de Investigación (PI) que se presenta en este trabajo, estuvo dirigido a desarrollar productos Web, aplicando y validando propuestas integradoras de técnicas y herramientas conceptuales y prácticas basadas en la eXperiencia de Usuario (UX), a los efectos de satisfacer las necesidades y preferencias de grupos de usuarios de interés en el diseño y desarrollo de productos multi-dispositivos.

Palabras clave: *Experiencia de Usuario (UX) | Diseño y Evaluación | Usabilidad | Accesibilidad Web | Grupos de Usuarios | Diseño de Experiencia Multi-Dispositivos.*

CONTEXTO

Nuestro 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 ejecutado en 2018 el primer año del Proyecto de Investigación (PI) Nº29/B222, Período: 2018-2020, denominado: "*Diseño y Evaluación de Experiencia de Usuario para Multi-Dispositivos*", dirigido por la Dra. Martín y codirigido por la Mg. Gaetán. Para desarrollar la problemática propuesta por este PI, GIFIS cuenta con sólidos antecedentes recabados a partir de la ejecución de los siguientes proyectos: PI Nº29/B194 (2016-2018), denominado: "*Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web*"; PI Nº29/B167 (2014-2016), denominado: "*Identificación, Desarrollo y Uso de Soluciones Web Centradas en el Usuario*" y, PI Nº29/B144 (2012-2014), denominado: "*Diseño y Evaluación de Portales Web*".

En este contexto, en el PI Nº 29/B222, dirigimos nuestros esfuerzos a desarrollar productos multi-dispositivos aplicando técnicas y herramientas basadas en la experiencia del usuario (UX), que satisfagan los requerimientos de usuarios de interés.

1. INTRODUCCIÓN

Sin lugar a dudas, estamos inmersos en un mundo de experiencias multi-dispositivo [1]. Muchos aspectos de la vida cotidiana de las personas, implican interactuar horas con computadoras de escritorio o portátiles, teléfonos inteligentes, *tablets*, televisores, relojes y brazaletes multi-propósito, consolas de

videojuegos, entre otros. La interacción multi-dispositivo se ha convertido en parte del día a día de nuestras vidas, ya sea para trabajar, estudiar, administrar cuentas bancarias y ciudadanas, etc., como así también, seguir programas favoritos, realizar ejercitación, participar activamente de redes sociales o simplemente navegar para crear espacios de recreación. Muchas personas, pasan más tiempo interactuando con dispositivos que con otras personas y, muy a menudo, lo hacen utilizando más de un tipo de dispositivo a la vez [1]. Es muy común, que las personas posean múltiples dispositivos conectados a sus computadoras personales y laborales, y los utilicen en conjunto para desplegar sus tareas y alcanzar sus objetivos. Según [2], en 2012, el número de dispositivos conectados en todo el mundo alcanzó 8.700 millones, y las previsiones para 2020 de las empresas líderes en el sector, arrojan cifras superiores a los 50.000 millones. Estos números, dan testimonio indiscutible de la preponderancia absoluta que tienen y seguirán teniendo los dispositivos en nuestra vida digital cotidiana. Ahora bien, en este escenario y tal como se señala en [1], una pregunta clave a responder es: ¿Qué significa diseñar un producto en un mundo donde las personas poseen varios dispositivos y los utilizan combinados y de forma intercambiable? Diseñar un producto para una red de redes a la que acceden e interactúan miles de millones de usuarios, todos ellos representativos de la diversidad humana, no es un tema menor. Desde el punto de vista del diseño, en general, la mayoría de los productos ofrecen un modelo único y completo, que independientemente del dispositivo seleccionado por el usuario para el despliegue, se caracteriza por: (i) estar basado en la experiencia de computadoras de escritorio; (ii) ofrecer siempre todo el contenido y funcionalidad; e (iii) incluir sólo algunos ajustes menores para adaptarlo al tamaño de los dispositivos. Este modelo genérico responde a una filosofía de diseño "*one-model-fits-all*", o como también señalan otros autores [1] [3], a un enfoque "*consistent design*". Si bien, esta forma de pensar al diseñar es adecuada, valiosa y reconocida, no toma en consideración las preferencias, comportamiento y necesidades de uso de las personas usuarias de multi-dispositivos. Estos factores son claves para la mejora de la UX en el acceso e interacción con productos desde diferentes tipos de dispositivos.

Desarrollar productos con la mente puesta en satisfacer al usuario de esos productos, no es una problemática nueva, y existe mucho trabajo previo y en curso, dedicado a proponer asistencia al diseño desde distintas perspectivas. Por ejemplo: (i) desde las perspectivas de cómo tratar la apariencia y distribución, el contenido y la información [4][5][6][7][8][9]; (ii) desde la perspectiva de asegurar la presencia de propiedades y atributos de calidad [10][11]; (iii) desde la perspectiva de aplicar enfoques para atender a la UX [1][12][13]; etc.

Tal como ya señalamos, GIFIS ha estado trabajando fuertemente en el marco de varios PI-UNPA, lo que constituye un sólido antecedente y motivación para definir los Objetivos del PI Nº29/B222:

Objetivo General: "*Desarrollar productos multi-dispositivos aplicando técnicas y herramientas basadas en la experiencia del usuario (UX), que satisfagan los requerimientos de grupos de usuarios de interés.*"

Objetivos Específicos (para usuarios de interés):

Objetivo Específico 1. Seleccionar técnicas y herramientas para desarrollar productos multi-dispositivos.

Objetivo Específico 2. Definir requerimientos de experiencias multi-dispositivos.

Objetivo Específico 3. Diseñar experiencias multi-dispositivos.

Objetivo Específico 4. Validar experiencias multi-dispositivos.

Para alcanzar el Objetivo General, el PI Nº29/B222 ha definido un Plan de 4 Actividades. Cada Actividad se corresponde con un Objetivo Específico y se compone de un conjunto de Tareas a ejecutar:

Actividad 1 → Objetivo Específico 1

Tarea1.1 Evaluar patrones y estrategias de diseño multi-dispositivo.

Tarea1.2 Analizar y seleccionar herramientas de conceptualización y visualización de UX.

Tarea1.3 Evaluar y seleccionar herramientas para prototipado y evaluación de UX.

Tarea1.4 Seleccionar criterios para establecer una estrategia de diseño.

Entregables: Guía de técnicas y herramientas recomendadas.

Actividad 2 → Objetivo Específico 2

Tarea2.1 Analizar y entender las necesidades de usuarios de interés.

Tarea2.2 Identificar patrones de comportamiento y tendencias de uso de usuarios de interés.

Tarea2.3 Identificar sinergias derivadas del uso combinado de varios dispositivos.

Entregables: Fichas de persona, Historias de usuario, Escenarios, Resultados de evaluación de Accesibilidad y Usabilidad.

Actividad 3 → Objetivo Específico 3

Tarea3.1 Analizar los diferentes modos de interacción disponibles en los dispositivos para los que se diseñarán las experiencias de usuario.

Tarea3.2 Desarrollar el diseño de IU a través de prototipos de baja y alta fidelidad.

Tarea3.3 Definir la estructura de navegación del producto.

Entregables: Prototipos, Mapas de sitio, Esquemas de página, Arquitectura de Información, Estrategia de contenidos

Actividad 4 → Objetivo Específico 4

Tarea4.1 Revisión de prototipos con usuarios de interés.

Tarea4.2 Evaluación de resultados e incorporación de mejoras.

Entregables: Informe de evaluación de UX.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Durante 2018, GIFIS ha estado enfocado en la problemática de mejorar la UX cuando estos interactúan con aplicaciones desplegadas en multi-dispositivos. A continuación, se describen las Líneas de Investigación (LI) basadas en usuarios pertenecientes a un determinado grupo y/o dominio de interés. Cabe señalar que los usuarios son el centro para el desarrollo del Plan de Actividades:

LI.1: Usuarios Adultos Mayores

Desde 2013, GIFIS ha estado vinculado al dictado de los cursos de computación en el marco del convenio UPAMI. A partir de 2016, GIFIS está a cargo de la elaboración de la propuesta y dictado de los cursos a los abuelos de la región. Este espacio de intercambio e interacción ofrece el marco adecuado para ejecutar la LI.1, la cual durante el período 2017-2018, ha estado enfocada en mejorar la UX de los Adultos Mayores (AM) con las redes sociales y desde dispositivo del tipo *tablet*.

LI.2 Usuarios del Dominio Universitario

Desde 2013, GIFIS ha estado trabajando sobre diferentes aspectos de diseño de los sitios Web universitarios en Argentina, para mejorar la interacción y acceso de los usuarios de la comunidad. Desde 2016, y basados en la interacción con estos sitios desde dispositivos móviles, se ha estado trabajando en propuestas [14][15], destinadas a ofrecer una mejora en la UX. Esta creciente demanda de satisfacer las nuevas necesidades de la comunidad universitaria, ofrece el marco adecuado para ejecutar la LI.2.

LI.3 Usuarios de Gobierno Electrónico

Desde 2017, GIFIS ha iniciado un estudio de los sitios Web gubernamentales, para analizar la UX de los ciudadanos con las aplicaciones de gobierno móvil en Argentina. Este estudio ofrece el marco adecuado para ejecutar la LI.3.

3. RESULTADOS OBTENIDOS/ ESPERADOS

En lo que se refiere a resultados obtenidos, y tan sólo considerando el período 2016-2018, podemos citar las siguientes contribuciones [14][15][16][17][18][19][20][21][22][23][24][25][26][27][28][29], entre otras.

3.1. Resultados Esperados

Alineados al Objetivo General-Objetivos Específicos y las Líneas de Investigación (Sección 1. y 2., respectivamente), a la culminación del PI Nº29/B222, se espera obtener:

- Enfoques que permitan el diseño y la evaluación de la UX y,
- Propuestas móviles concretas que permitan validar estos enfoques, al demostrar mejoras en la UX de usuarios de interés.

3.2. Resultados Obtenidos

Durante 2018, las contribuciones C-T vinculadas a las Líneas de Investigación del PI 29/B222, son las siguientes:

- En la LI.1: Usuarios Adultos Mayores, se trabajó fuertemente con las redes sociales accedidas desde dispositivos de tipo *tablet*, para mejorar la UX de los AM [21][26][29].

- En la LI.2: Usuarios del Dominio Universitario, se trabajó en mejorar la UX aplicando estrategias de contenido en un sitio Web móvil, como así también, técnicas de diseño en el desarrollo de una aplicación móvil de *Carpooling* [25] y [22][28], respectivamente.
- En la LI.3: Usuarios de Gobierno Electrónico, se analizó la UX en las aplicaciones de gobierno móvil [27].

Además, los integrantes del GIFIS estuvieron presentes con contribuciones C-T en los siguientes eventos 2018:

- XX WICC, se envió 1 artículo [23] y su respectivo poster.
- 5º EIPA, se presentaron 5 artículos [24][25][26][27][28] y se viajó a realizar las ponencias con los respectivos posters.
- 6to. CONAIISI, se presentó el 1 artículo [29] y se viajó a realizar la exposición del mismo.

Otros resultados 2018 vinculados el PI 29/B222, son la ejecución de las siguientes actividades de extensión:

- Programa de Extensión y Vinculación (PgEyV): "Observatorio de Experiencia de Usuario"; director Mg. Gaetán; codirector Mg. Saldaño.
- Proyecto de Extensión y Vinculación 1 (PEyV1): "Taller para Adultos Mayores: ¿Cómo usar los Nuevos Teléfonos Móviles?"; director y docente responsable Mg Gaetán; equipo de trabajo Mg. Saldaño y Dra. Martín.
- Proyecto de Extensión y Vinculación 2 (PEyV2): "Abuelos del Siglo XXI: Taller de Redes Sociales para Adultos Mayores"; director y docente responsable Mg Saldaño; equipo de trabajo Mg. Gaetán y Dra. Martín.

A partir de la experiencia recabada por los integrantes de GIFIS en la elaboración de ICT-UNPA, la Dra. Martín, Mg. Saldaño y Mg. Gaetán, dictaron un curso de postgrado y extensión, denominado: "Consignas para la Elaboración de ICT-UNPA".

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 29/B222 cuenta con 3 integrantes alumnas de postgrado, las cuales están cursando la Maestría en Informática y Sistemas (MIS-UNPA). Durante 2018, estas alumnas tuvieron los siguientes avances:

- 1 maestranda terminó de escribir su tesis MIS-UNPA, la cual entregó a revisión en febrero de 2019.
- 2 maestrandas están escribiendo sus respectivos ICT-UNPA requisito de la tutoría MIS-UNPA.

5. AGRADECIMIENTOS

A la UNPA por dar soporte al PI Nº: 29/B222, Período: 2016-2018, denominado: "*Diseño y Evaluación de Experiencia de Usuario para Multi-Dispositivos*".

6. REFERENCIAS

[1] Levin, M. Designing Multi-Device Experiences: An Ecosystem Approach to User Experiences Across Devices. Ed: O'Reilly; 2014. ISBN: 978-1-449-34038-4.

[2] Statista. El Portal de Estadísticas; disponible-en: <https://es.statista.com/estadisticas/638100/internet-de-las-cosas-numero-de-dispositivos-conectados-en-todo-el-mundo--2020/>; accedido Marzo 2019.

[3] Nikolov, A. Design principle: Consistency: The most known and the most fragile design principle. UXdesign.cc; disponible-en: <https://uxdesign.cc/design-principle-consistency- 6b0cf7e7339f>; accedido: Marzo 2019.

[4] Krug, S."Don't Make Me Think". Pearson Education; 2009.

[5] McGrane, K. "Content Strategy for Mobile". A Book Apart; 2012.

[6] Morville, P. "Information Architecture". O'Reilly Media, Digital Edition; 2002.

[7] Norman, D. "Design of Everyday Things". Newprint, Digital Edition; 2002.

[8] Redish, J. G. "Letting Go of the Words: Writing Web Content that Works". Morgan Kaufmann, 2nd Edition; 2012.

[9] Tidwell, J. "Designing Interfaces", O'Reilly Media; 2nd Edition; 2009.

[10] Nielsen, J. "Usability Engineering". Elsevier Science; 1994.

[11] W3C-WAI. Web Content Accessibility Guidelines (WCAG); disponible-en: <https://www.w3.org/WAI/intro/wcag>; accedido: Marzo 2019.
[12] Garrett, J. J. "Elements of User Experience: The User-Centered Design for the Web and Beyond". Pearson Education; 2010.
[13] Hartson, R.& Pyla, P. S. "The UX Book: Process and Guidelines for Ensuring a Quality User Experience". Elsevier; 2012.
[14] Moyano, A., Gaetán, G., Martín, A. Interfaz Móvil para el Sitio Web de la UACO. Un Prototipo centrado en el Usuario. ICT-UNPA 2016; vol. 8 (1); pp. 172-201; ISBN: 1852-4516; DOI: http://dx.doi.org/10.22305/ict-unpa.v8i1.156
[15] Gaetán, G., Martin, 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.
[16] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Miranda, G., Sosa, H. Un Enfoque Integrador para Diseñar y Evaluar Interfaces de Usuario Web. 4º EIPA 2016; Caleta Olivia.
[17] Cardozo, C., Saldaño, V., Martin, A., Gaetán, G. Los Adultos Mayores y la Utilización de Redes Sociales en Dispositivos Móviles. 4º EIPA 2016; Caleta Olivia.
[18] Miranda, G., Gaetán, G., Martin, A., Saldaño, V. Un Enfoque de Desarrollo Basado en la Experiencia de Usuario (UX) para favorecer la Accesibilidad Web. 4º EIPA 2016; Caleta Olivia.
[19] Cardozo, C., Martín, A., Saldaño, V. Los Adultos Mayores y las Redes Sociales: Analizando Experiencias para Mejorar la Interacción. ICT-UNPA 2017, vol. 9(2), pp. 1-29. ISSN: 1852-4516; DOI: http://dx.doi.org/10.22305/ict-unpa.v9i2.244
[20] 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. CONAIISI 2017, 2-3 Noviembre; Santa Fé; Argentina; pp. 478-487; ISSN: 2347-0372.
[21] Cardozo, C., Martín A., Saldaño, V. Recomendaciones de Diseño para Mejorar la Experiencia de los Usuarios Adultos Mayores con Facebook en Dispositivos Tablet. ICT-UNPA 2018; vol. 10(2); pp. 1-32; ISSN: 1852-4516; DOI: http://dx.doi.org/10.22305/ict-unpa.v10i2.268.
[22] Vargas, F., Gaetán, G.; Saldaño, V. Usando Personas para Mejorar la Experiencia de Usuario de una Aplicación Móvil de Carpooling. ICT-UNPA 2018; vol. 10(2); pp. 96-116; ISSN: 1852-4516; DOI: http://dx.doi.org/10.22305/ict-unpa.v10i2.273.
[23] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Carrizo, A., Vargas, F. Proponiendo Un Enfoque Integrador Para Diseñar y Evaluar Interfaces de Usuario Web. WICC 2018; Corrientes.
[24] Martin, A., Gaetán, G., Saldaño, V., Cardozo, C., Villagra, S., Carrizo, A., Vargas, F. Enfoques para el Diseño y Evaluación de Experiencia de Usuario en Multi-Dispositivos. 5º EIPA 2018; Río Gallegos.
[25] Villagra, S., Gaetán, G., Saldaño, V., Martin, A. Ecosistema de Contenido para un Sitio Web Móvil Universitario. 5º EIPA 2018; Río Gallegos.
[26] Cardozo, C., Martin, A., Saldaño, V., Gaetán, G. Adultos Mayores y Redes Sociales: Una Propuesta de Diseño para Mejorar la Experiencia en Dispositivos Tablet. 5º EIPA 2018; Río Gallegos.
[27] Carrizo, A., Gaetán, G., Saldaño, V., Martin, A. Análisis de la Experiencia de Usuario con Aplicaciones de Gobierno Móvil en Argentina. 5º EIPA 2018; Río Gallegos.
[28] Vargas, F., Gaetán, G., Saldaño, V., Martin, A. Mejoras en la Experiencia de Usuario para una Aplicación Móvil de Carpooling. 5º EIPA 2018; Río Gallegos.
[29] Cardozo, C., Martin, A., Saldaño, V., Gaetán, G. Una Propuesta de Red Social para Dispositivo Tablet: Mejorando la Experiencia de los Adultos Mayores. 6to. CONAIISI 2018, 29-30 Noviembre; Mar del Plata; Argentina (actas en elaboración).

EL DISEÑO Y LAS EMOCIONES EN LA EXPERIENCIA DEL USUARIO

Eugenia Sosa Bruchmann[1]; Chayle, Carolina I. [2], Germán Montejano[3]; Ana Garis[3] Martin, Luis E. [1]
[1]Departamento Computación/Facultad de Ciencias Exactas y Naturales/Universidad Nacional de Catamarca

Av. Belgrano Nº 300 - Planta alta - C.P: 4700 - San Fernando del Valle de Catamarca
Teléfono: 0383- 420900
[2] Departamento Informática/Facultad de Tecnología y Ciencias Aplicadas /Universidad Nacional de Catamarca

[3] Facultad de Cs. Físico-Matemáticas y Naturales/Universidad Nacional de San Luis
Avenida Ejército de Los Andes 950
Teléfono: 0266- 4424027
sosab_ec@hotmail, cchayle@gmail.com

Resumen

Es importante precisar a la Experiencia de Usuario UX, como una filosofía de diseño, cuyo objetivo es la creación de sistemas interactivos que brinden soluciones inmediatas por medio de experiencias únicas y satisfactorias para el usuario. Para ello, se debe llevar un paso más allá a la arquitectura de la información de un sitio, considerando no sólo su navegación, sino también la habilidad del usuario para realizar una acción, con el fin de facilitarle esta tarea. Básicamente, el diseño desde el punto de vista de UX consiste en trabajar para hacer cosas más profundas, realizando una clasificación de los usuarios desde un nivel emocional. Por consiguiente, se ha podido determinar como un conocimiento sustancial a resaltar durante estas primeras aproximaciones a la investigación para este proyecto, un nuevo concepto aun no muy difundido, que podría resultar clave para conseguir los resultados esperados, la deseabilidad, la cual se puede especificar por el grado en que un producto mejora la vida de las personas y las hace más felices. Es aquí en donde se añade este plus, que consiste en determinar la satisfacción pero desde el punto de vista de la emoción, que es la esencia de la investigación de este proyecto.

Palabras clave: Experiencia de Usuario; emoción, Satisfacción de uso.

Contexto

Este trabajo representa los avances realizados, dentro del Proyecto Anual que se encuentra en desarrollo, enmarcado en el Programa de Desarrollo Científico y Tecnológico de la Secretaria de Ciencia y Tecnología de la Universidad Nacional de Catamarca, denominado: El Diseño y las Emociones en la Experiencia del Usuario.

Introducción

Es sabido que los proveedores de artefactos Web, escasamente ofrecen a sus clientes diseños que garanticen que la metodología del producto o servicio que vayan a vender, se ajuste a estándares de Usabilidad y que se

adapten a las características del Diseño Centrados en el Usuario. Ante esta realidad, se percibe que los diseñadores deben comprender profundamente a sus Usuarios específicos para adaptar a ellos, los contenidos. Sin embargo, dada la complejidad del ser humano en la Interacción Persona-Ordenador, resulta casi imposible la creación de un diseño que sea adecuado, en un primer intento. Es por ello que entender a una audiencia y a un Usuario, requiere de una investigación y pensamiento crítico. Podría decirse, saber "concebir al Usuario". Asimismo, los diversos modos de aprendizaje, también reflejan variantes que se manifiestan continuamente. Existen ligeras preferencias además de profundas necesidades en cada uno de los Usuarios. Por lo tanto, se debe amoldar esta diversidad, mediante representaciones alternativas de la información que sea clave, ya que la organización de la información de un sitio Web va a resultar más útil, cuanto más adaptada se encuentre al modelo mental de sus Usuarios. Las preferencias y necesidades particulares de cada uno de ellos, siempre encontrarán medios, sustentos y opciones que permitan al mismo mostrar su conocimiento, de la forma que le resulte más efectiva.

Aparte de la teoría cognitiva, ¿Existe alguna otra perspectiva teórica, que se establezca como una representación más integradora y a su vez más útil, al momento de explicar las emociones provocadas por los productos o Sitios Web? Una de las cuestiones más relevantes de esta investigación será corroborar que la Experiencia de Usuario se verá influenciada tanto por el comportamiento racional o cognitivo del usuario, como por su comportamiento emocional o afectivo y, aunque estén interrelacionados, influirán de diferente modo en la llamada Satisfacción del Usuario [1,2].

Es así que, para tales efectos se consideran pertinentes realizar un estudio y análisis profundo, de los denominados Modelos Cognitivos, que son los que recogen la Experiencia Emotiva del Usuario [3].

Líneas de Investigación, Desarrollo e innovación

Se realizara el análisis de diferentes Modelos Cognitivos a fin de determinar los aportes sobre las ventajas de un sitio Web estéticamente agradable, a fin de determinar la percepción al mismo tiempo por parte del usuario, como el "más fácil de usar". Es decir determinar cuál es la relación que existe entre la estética y la satisfacción final del Usuario, si estas variables se hayan correlacionadas o no, además de cuándo o bajo qué condiciones se produce o no esta correlación [4].

Las técnicas que se aplicaran en este proyecto se encuentran enmarcadas en el denominado Diseño Centrado en el Usuario [5,6].

- Entrevistas y Encuestas, de tal manera de tener un contacto personalizado con los usuarios. Luego la información obtenida será analizada de manera cualitativa y cuantitativa. Esta es una de las metodologías más concretas para asegurar que la Experiencia de Usuario se está resguardando adecuadamente.
- Revisión de Productos o Artefactos Web similares entre los cuales se seleccionaran entre otros, sitios educativos de la FaCEN, a quienes se les realizara, un análisis de similitud, el cual se sustentara en aspectos de contenido, diseño y/o programación.
- Definición de cuales indicadores serán evaluados, a saber:
    - Niveles de navegación.

    - Mantenimiento de la imagen de diseño en todo el producto.
    - Buena funcionalidad, etc.
- Evaluación de indicadores en cada producto escogido y se tabularan los resultados para una mejor comprensión de los mismos.

Resultados y Objetivos

Existen algunos trabajos sobre la temática y resultados de esta y diferentes investigaciones, que han sido producidos por el Director. Se pretende realizar asimismo, por parte del equipo de investigación una exploración de las distintas perspectivas en torno a los vínculos entre afectividad y diseño. Además, se proyecta que todos los integrantes presenten resultados de esta investigación en los cuales se profundicen los conceptos de las emociones evocadas, la empatía y satisfacción de uso, contribuyendo así a la incorporación de la temática al estado del arte del tema investigado.

Los objetivos planteados en la Investigación son:

- Conocer acabadamente el método de Diseño Centrado en el Usuario, las prácticas de investigación y evaluación con usuarios verdaderos, el diseño de la Interacción y la Arquitectura de la información.
- Identificar y analizar las características de sitios evaluados por usuarios, a fin de determinar los factores que influyen sobre la probabilidad de que los mismos realicen nuevas visitas.
- Adaptar y habituarse a las reglas Heurísticas, así como ejercitar las buenas prácticas de diseño.
- Disponer una aproximación teórica de lo que significa la Satisfacción de Uso de sitios Web, realizando un aporte de un modelo explicativo desde la perspectiva del impacto de los diferentes factores de diseño en dicha satisfacción.
- Establecer que variables intervienen principalmente en el comportamiento cognitivo del usuario durante la interacción, y qué papel desempeña la estética como atributo que interviene principalmente en el comportamiento afectivo-emocional del usuario, de tal manera de determinar cómo estos atributos y comportamientos interactivos influyen en la satisfacción de uso del sitio.
- Conocer que la Usabilidad percibida presenta una alta correlación con otros atributos del diseño, ajenos a la Usabilidad objetiva o inherente, como es la estética o apariencia visual.

Cabe destacar que la temática de este proyecto es considerada un Área de Vacancia dentro de la Facultad de Ciencias Exactas y Naturales de la UNCa, por lo que resultan interesantes y enriquecedores los resultados obtenidos en esta primera etapa del Proyecto, se estima se seguirá investigando y ahondando sobre la temática, con mayor profundidad en futuros proyectos de investigación.

Formación de Recursos Humanos

El equipo de investigación se encuentra compuesto por profesionales que tienen una vasta experiencia en el ámbito investigativo, han participado de sólidos proyectos de Investigación financiados por la Secretaria de Políticas Universitarias y la Secretaria de Ciencia y Tecnología de la UNCa, como Integrantes. Así como de Proyectos Internacionales. Además son integrantes de Centros de Investigación en la Facultad de Ciencias Exactas y Naturales y Facultad de

Tecnología y Ciencias Aplicadas de la UNCa. Desde el Proyecto de investigación se propone cimentar en sus trabajos de tesis de postgrado. Tal es el caso de los tres integrantes quienes desarrollan su tesis de Maestría en la Temática, y quienes son de gran ayuda y brindan importante colaboración en los aspectos investigativos de la Temática del proyecto. Se realizaron y realizan actualmente cursos de postgrado como así también participan activamente en eventos científicos que permitirán profundizar sus conocimientos en la temática de investigación.

Referencias

[1] Arhippanien y Tahti. (2003). *"Empirical Evaluation of User Experience in Two Adaptive Mobile Application Prototype"*

[2] Bagozzi, R., (2007). *"The legacy of the Technology Acceptance Model and a Proposal for a Paradigm "* Shift. Journal of the Association for Information System, 8 (4), 244-254.

[3] Granollers, T.; Lorés, J. (2003). *"Tendencias Actuales en la interacción Persona Ordenador: Accesibilidad, Adaptabilidad y Nuevos Paradigmas"*.(ISBN: 84-921873-1-X). Capítulo: La Ingeniería de la Usabilidad aplicada al diseño y desarrollo de sitios web, pp 119-144. Ediciones de la Universidad Castilla-La Mancha (Albacete, España); XIII Escuela de Verano de Informática, julio 2003.

[4] Mendez, Y. A., Collazos, C. A., Granollers, T., & Sanchez, J. A. (2014). *"Evaluating interactive systems from an emotional perspective"*. Revista Científica *Guillermo de Ockham, 12(1), 43-49.

[5] Hassenzahl, M. (2011). Encyclopedia chapter titled *"User Experience and Experience* Design". Retrieved 14 June 2011 from Interaction-Design.org.

[6] Covella, G. J. (2005). "Medición y Evaluación de Calidad en Uso de Aplicaciones *Web"*. Grupo de Investigación y Desarrollo en Ingeniería de Software e Ingeniería Web (GIDISWeb), Facultad de Ingeniería de la UNLPam.

[7] Sosa Bruchmann, E.; Montejano, G.; Garis, A. (2015). Análisis de UX: Relación entre el Comportamiento Emocional y la Satisfacción de Uso.

[8] Desmet, P.M.A. (2004). *"From disgust to desire; how products elicit emotions."* In D.C. McDonagh, P. Hekkert, J. van Erp, and Gyi (Eds.), Design and Emotion: the experience of *everyday things, pp. 8-12. London: Taylor & Francis.*

[9] Hassan-Montero, Y. (2006). "Factores Del Diseño Web Orientado a la Satisfacción y No-Frustración de Uso". Revista Española de Documentación Científica - 29, 2, Abril-Junio, pp. 239-257, 2006. ISSN 0210-0614. Disponible en: http://www.yusef.es/publicaciones.htm.

[10] Hassan-Montero, Y.; Ortega-Santamaría, S. (2009). *"Informe APEI sobre Usabilidad"*. Gijón: Asociación Profesional de Especialistas en Información, 2009, pp. 73. ISBN: 978-84-692-3782-3.

[11] Hassenzahl, M., Burmester, M. & Koller, F. (2008). *"Der User Experience (UX) auf der Spur: Zum Einsatz von"* www.attrakdiff.de. In: Brau, H., Diefenbach, S., Hassenzahl, M., Koller,F., Peissner, M. & Rose, K. (Hrsg.): Usability Professionals 2008, S. pp.78-82. Stuttgart: German Chapter der Usability Professionals Association

[12] Hassenzahl, M., & Monk, A. (2010). "The inference of Perceived Usability from Beauty ". Human-Computer Interaction, 25 (3), pp. 235-260.

[13] Hassenzahl, M. (2011). Encyclopedia chapter titled " User E*xperience and Experience* Design". Retrieved 14 June 2011 from Interaction-Design.org. Disponible en: http://www.interactiondesign.org/encyclopedia/user_experience_and_experience_design.html.

[14] Keinonen, T. (2003). Publicado originalmente como Capítulo 2 de: "One-dimensional usability - influence of usability on consumers' product preference" A21 de la UIAH. Helsinki 1998. Disponible en: http://www2.uiah.fi/projects/metodi/

[15] Lindgaard, G. & Millard, N. (2002). "The business value of HCI: How can we do better?", Proceedings 20th. Conference on Human Factors in computing systems, CHI'02, pp 928-929. Disponible en: http://dl.acm.org/citation.cfm?id=506663

[16] Norman, D., (2004). "Emotional Design: Why We Love (or Hate) Everyday Things". Basic Books, 224 pp. Book revieus Ramarao Desiraju Editor University of Central Florida.

[17] Petrie, H. (2010). "Measuring User Experience of Websites: Think aloud protocols and an emotion word prompt list Methods"- 3673-3678. doi:10.1145/1753846.1754037.

[18] Spillers, F.(2013). "Emotion as a Cognitive Artifact and the Design Implications for Products That are Perceived As Pleasurable". Recuperado el 18 de febrero de 2013 en:

[19] Travis, D. (2002). "E-commerce Usability-2002- Tools and techniques to perfect the on-line experience". Disponible en: http://www.userfocus.co.uk/book/about/index.html Disponible en: <http://www.interactiondesign.org/encyclopedia/user_experience_and_experience_design.html>.

[20] Desmet, P., & Hekkert, P. (2007b). "Framework of Product Experience". International Journal of Design, 1(1), pp. 57-66.

Evaluación de herramientas libres para la validación de requerimientos

Sonia R. Santana[1], Lucrecia R. Perero[1], Amalia G. Delduca[1],
Gladys N. Dapozo[2]
(1) Facultad de Ciencias de la Administración - Universidad Nacional de Entre Ríos
(2) Facultad de Ciencias Exactas y Naturales y Agrimensura - Universidad Nacional del Nordeste
e-mail: sr.santana675@gmail.com

Resumen

La validación de los requerimientos es una tarea fundamental en cualquier proyecto de Ingeniería de Software. Debe ser un proceso continuo en el ciclo de vida del desarrollo del sistema. Los requerimientos son volátiles debido a los constantes cambios en las necesidades, procesos y tecnologías, esto hace que la gestión de los requerimientos en forma manual sea una tarea difícil. Existen herramientas software que facilitan la gestión de requerimientos. En este proyecto se abordará el estudio y profundización de herramientas de software libre orientadas a la validación de requerimientos. El software libre presenta características y modos de usos que lo hacen accesible y apropiado para esta tarea. Se realizará un análisis comparativo de las herramientas para determinar la performance de su aplicación en base a determinados criterios de evaluación establecidos previamente.

Palabras clave: Ingeniería de Requerimientos, validación de requerimientos, herramientas para la ingeniería de requerimientos, software libre.

Contexto

El presente PID 7057 se encuadra en la línea de investigación "Ingeniería de Software", establecida como prioritaria desde la carrera Licenciatura en Sistemas de la Facultad de Ciencias de la Administración de la Universidad Nacional de Entre Ríos (UNER). Se adecua, además, a una de las prioridades de la UNER considerando que es un proyecto aplicado a la investigación sobre Tecnologías de la Información y la Comunicación.

Introducción

El desarrollo de software no es solo un desafío tecnológico sino también un proceso social complejo. En este proceso, la comunicación efectiva entre los interesados desempeña un papel vital en el desarrollo del proyecto de software que conduce a recopilar requerimientos de manera correcta y consistente.

Los errores de requerimientos tales como inconsistencia, incompletitud e incorrección pueden llevar al retrabajo extenso y fallas irrecuperables [1]. Además, la detección de estos errores en una etapa posterior de un proyecto de desarrollo de software es más difícil, lento y costoso [2] [3].

Es importante lograr que los errores se detecten, se corrijan y minimicen lo antes posible. La prevención de errores es una cuestión de buenas prácticas en la Ingeniería de Software, sin embargo, es prudente suponer que se producirán errores. Afortunadamente, este tipo de errores innecesarios pueden evitarse con una adecuada validación de requerimientos.

El principal objetivo de la validación de los requerimientos es certificar que los requerimientos elicitados son representaciones exactas de las necesidades y expectativas de los usuarios [4] [5] [10] y deben ser completos, correctos y consistentes [11].

Esta actividad clave puede ayudar a identificar, prevenir los defectos y errores, mejorando la calidad de los requerimientos, reduciendo el tiempo de desarrollo, el costo y los riesgos para el desarrollo del software de alta calidad que cumpla con las expectativas de los usuarios [2] [6].

La validación de requerimientos debe considerarse más como una actividad de "debugging" en lugar de probar "lo correcto". En efecto, es una actividad "siempre presente" que tiene lugar cada vez que se solicita que un requerimiento sea adquirido, analizado o integrado con el resto del modelo de requerimientos [7].

La investigación en el área de Ingeniería de Requerimientos (ER) ha sido reconocida desde mediados de la década del 80. Desde su creación, ha habido muchas literaturas o estudios de mapeo sistemáticas que cubren varios aspectos de ER tales como elicitación, jerarquización y análisis de requerimientos.

Diferentes revisiones sistemáticas publicadas demuestran el interés en la investigación empírica de la ER que está en aumento en su conjunto. Sin embargo, la evaluación de herramientas para la validación de los requerimientos, a pesar de su reconocida importancia, la investigación empírica es todavía incipiente en la actualidad [8].

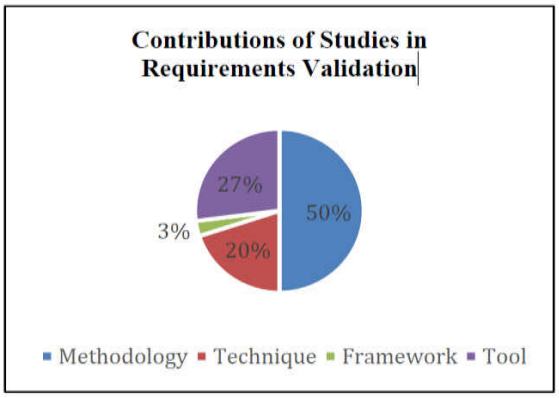


Figura 1. Contribución de estudios en la validación de requerimientos desde el año 2007 hasta el año 2016 [8].

La gestión de requerimientos es fundamental para el éxito de entrega de software y ciclo de vida del proyecto. Proporciona asistencia para muchas actividades de ingeniería de software como análisis de impacto, análisis de responsabilidad, validación de requerimientos y pruebas de repercusión.

Se define gestión de requerimientos como: "Proceso de administrar los cambios a los requerimientos para asegurarse de que los cambios realizados se analizan adecuadamente y se rastrean a lo largo del sistema" [10].

Las herramientas de ER forman parte integral de la mayoría de las soluciones de gestión de requerimientos. Son necesarias por varias razones, permiten:

- administrar un número creciente de información,
- mantener la trazabilidad entre los requerimientos, la solución y las pruebas,
- evaluar el impacto de un cambio de requerimientos en la solución y las pruebas.

Las herramientas de ER se utilizan cada vez más para facilitar los procesos de ER y permitir un manejo más sistemático y formal de los requerimientos, la gestión de cambios y la trazabilidad.

Muchos investigadores han desarrollado y estudiado varias herramientas que ayudan a los profesionales en el proceso de ER Algunos de los estudios se dan de la siguiente manera:
Atif Shah [15] realizó un análisis comparativo de varias herramientas que fueron evaluadas por su funcionalidad, alcance y orígenes geográficos. Según el análisis muestra que la mayoría de las herramientas no son compatibles con la trazabilidad, la representación gráfica y la integración con otras herramientas. El alcance determinó que todas las herramientas cumplen con los procesos de ER. Las diversas herramientas tratadas se desarrollan en diferentes ubicaciones geográficas.
Carrillo de Gea [9] evaluaron herramientas de RE, basados en encuestas y experiencias de usuarios, donde señalan que no apoyan por completo la gestión de requerimientos. A pesar de que manejan bien la obtención de requerimientos, pocos de ellos son compatibles con editores de requerimientos para llenar el vacío que implica compartir, comunicarse y colaborar de manera efectiva a través de diferentes herramientas. El soporte para editores de requerimientos es necesario porque las empresas generalmente no trabajan en el mismo repositorio de requerimientos ni utilizan las mismas herramientas de ER.

Hoy, las herramientas están evolucionando rápidamente. La demanda de flexibilidad, desarrollo ágil, colaboración mundial y el software avanzado están cambiando la manera en que se gestionan los requerimientos. Por ejemplo, los equipos ágiles están menos centrados en el documento y más orientados a los códigos, esperan requerimientos breves directamente relacionados con los cambios de código, por lo que su herramienta de ER debe ser liviana. Por otro lado, desarrollo distribuido los equipos necesitan acceder fácil y ampliamente a los requerimientos y las especificaciones con trazabilidad a lo largo de todo el ciclo de vida.

Las herramientas ER se están adaptando a estas demandas con cambios en el diseño y arquitectura. Tradicionalmente, orientadas a menudo a distintos entornos y nichos de mercado. Esto es motivo suficiente para evaluar las herramientas y tecnologías de ER con diferentes casos de uso [9].

En función de lo señalado, nuestro abordaje busca profundizar en herramientas para validar requerimientos disponibles en el ámbito del software libre.

Líneas de Investigación, Desarrollo e Innovación

En el marco de la investigación se pretende:

- Estudiar en profundidad los conceptos, métodos y técnicas para la gestión de requerimientos en el marco de la Ingeniería de Software.
- Relevar y seleccionar herramientas de software libre orientadas a la gestión de requerimientos, en especial, las aplicadas a la validación de los requerimientos.
- Evaluar las herramientas según criterios preestablecidos y obtener un cuadro resumen de la performance de las herramientas, según los aspectos considerados.

Resultados obtenidos/esperados

El análisis propuesto en este proyecto se sustenta en la necesidad de arribar a conclusiones generales y comparativas acerca de la performance de las herramientas libres para la validación de requerimientos y su relación con los distintos criterios de evaluación establecidos sobre las mismas.

El procedimiento metodológico adoptado para esta investigación se basa en tres etapas distintas.

En la primera etapa se seleccionarán diversas herramientas libres de RE teniendo en cuenta su funcionalidad y alcance, con el objetivo de establecer una contribución efectiva del presente proyecto frente al escenario actual.

Identificada las herramientas, en la segunda etapa se analizarán diferentes técnicas de validación existentes, tales como: checklist, revisiones de los requerimientos, el prototipado, los casos de prueba, los puntos de vista, animaciones y paráfrasis del lenguaje natural [12] [13] [14] [7], con el objetivo de seleccionar la técnica más adecuada para ser incorporada en el contexto.

Finalmente, como última etapa las herramientas serán analizadas con los siguientes criterios:

- Especificación de requerimientos.
- Prevención de defectos.
- Tiempo y esfuerzo.
- Ciclo de vida del software.
- Planificación.
- Procedimientos.
- Validación del software después de un cambio.
- Alcance de la validación.
- Independencia de la validación.
- Flexibilidad y responsabilidad.

Para formalizar la comparación entre los criterios de evaluación se orientará a responder cuestiones tales como:

a) Puntos fuertes y débiles de las herramientas libres en la validación de requerimientos.
b) Capacidad de las herramientas libres para asegurar la trazabilidad de los requerimientos desde su captura, modelado y seguimiento, es decir, que permita vincular los requerimientos a su fuente, a los cambios en los requerimientos y modelar elementos que cumplan con los requerimientos.
c) Disponibilidad de datos sobre la implementación de herramientas libres en la validación de sus requerimientos.

Con esta información sobre herramientas libres aplicadas a la validación de requerimientos se esperar contribuir con la actividad de los profesionales del sector de desarrollo de software. Además, la comparación presentada debería ayudar en el proceso de mejora continua de las herramientas existentes para la gestión de requerimientos.

Formación de Recursos Humanos

Este proyecto prevé la iniciación en actividades de investigación de tres docentes de la carrera Licenciatura en Sistemas, el desarrollo de, al menos, dos proyectos de Trabajo Final de la carrera Licenciatura en Sistemas y la realización de una tesis de maestría correspondiente a la Maestría de Ingeniería de Software de la Facultad de Informática de la UNLP.

Referencias

1. J. Zhou, "An Observer-Based Technique with Trace Links for Requirements Validation in Embedded Real-Time Systems," in International Conference on Research in Engineering and Technology, 2014, no. 177, pp. 1–52.
2. D. Aceituna, H. Do and S.-W. Lee, "SQ^ (2) E: An Approach to Requirements Validation with Scenario Question," in 2010 Asia Pacific Software Engineering Conference, 2010, pp. 33–42.
3. L. Kof, R. Gacitua, M. Rouncefield, and P. Sawyer, "Ontology and Model Alignment as a Means for Requirements Validation," in 2010 IEEE Fourth International Conference on Semantic Computing, 2010, pp. 46–51.
4. P. A. Laplante, Requirements Engineering for Software and Systems. CRC Press, 2009.
5. B. H. C. Cheng and J. M. Atlee, "Current and Future Research Directions in Requirements Engineering," in Design Requirements Engineering A Ten-Year Perspective. Lecture Notes in Business Information Processing, vol. 14, 2009, pp. 11–43.
6. M. Kamalrudin, N. A. Moketar, J. Grundy, and J. Hosking, "Automatic Acceptance Test Case Generation From Essential Use Cases," in 13th International Conference on Intelligent Software Methodologies, Tools and Techniques, 2014, pp. 246–255.
7. Loucopoulos, P., Karakostas, V., "System Requirements Engineering", McGraw-Hill, London, ISBN 0-07-707843-8, 1995.
8. Moketar, Nor Aiza & Kamalrudin, Massila & Mohd-Yusof, Mokhtar & Sidek, Safiah, "A review on requirements validation for software development", Journal of Theoretical and Applied Information Technology. 96. 3182-3193, 2018.
9. J. M. C. de Gea, J. Nicolas, J. L. F. Aleman, A. Toval, C. Ebert, and A. Vizcaino, "Requirements engineering tools," IEEE software, vol. 28, no. 4, pp. 86–91, 2011.
10. G. Kotonya and I. Sommerville, "Requirements Engineering: Processes and Techniques", Wiley, 1998.
11. S. L. Pfleeger, "Software Engineering – Theory and Practice", Prentice Hall, 1998.
12. I. Sommerville, "Ingeniería de Software", Novena Edición Pearson Educación, México, 2011.
13. C. Sandler, G. J. Myers, "The Art of Software Testing", 2nd ed., New York, Wiley, 2004.
14. J. C. S. P. Leite, P. A. Freeman, "Requirements Validation Through Viewpoint Resolution", IEEE Transactions on software Engineering, vol. 17, no.12, pp. 1253-1269, 1991.
15. Atif Shah, Mohamed Ali Alasow, Faisal Sajjad and Jawad Javed Akbar Baig "An Evaluation of Software Requirements Tools", 8th IEEE International Conference on Intelligent Computing and Information Systems, ICICIS 2017.

Evaluación Multicriterio Sobre Herramientas de Análisis de Seguridad en Aplicaciones Web

Ayrton Marini, Enrique A. Miranda, Mario Berón, Miguel Bustos, Daniel Riesco, Pedro R. Henriques

Departamento de Informática/ Universidad Nacional de San Luis/ Argentina

Departamento de Informática/ Universidade do Minho/ Portugal

mariniayrtond@gmail.com, {eamiranda,mberon,mabustos,driesco}@unsl.edu.ar, prh@di.uminho.pt

Resumen

En la actualidad, la mayoría de los sistemas informáticos utilizan Internet como su medio de despliegue. Un factor crítico de éxito durante su desarrollo, es garantizar su seguridad. Para llevar a cabo esta tarea, dentro de las alternativas más utilizadas, se encuadran las herramientas de análisis estático y dinámico de software. No obstante, existe una gran variedad de analizadores los cuales poseen diversas características, convirtiendo al proceso de selección de una herramienta que mejor se adecue a una determinada situación, una tarea tediosa para el ingeniero de software. Es por esto que es de vital importancia proveer mecanismos sistematizados y eficaces para evaluar las distintas alternativas.

El presente trabajo propone una línea de investigación que tiene dos objetivos relacionados. Por un lado, el estudio de diferentes métodos de evaluación multicriterio y por otra parte, su aplicación a la evaluación de herramientas de análisis de la seguridad de aplicaciones web.

Palabras clave: Seguridad Informática, Evaluación Multicriterio de Software, Herramientas de Análisis.

Contexto

La línea de investigación descripta en este artículo se encuentra enmarcada en el contexto de dos proyectos. Uno de ellos (P-031516) denominado: Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad de la Universidad Nacional de San Luis. Y otro (PO/16/93) llamado: Fortalecimiento de la Seguridad de los Sistemas de Software mediante el uso de Métodos, Técnicas y Herramientas de Ingeniería Reversa. Un proyecto bilateral internacional que surge de la relación entre la Universidad Nacional de San Luis y la Universidade do Minho oriunda de Portugal.

Introducción

Las aplicaciones web son utilizadas en una amplia variedad de áreas, entre ellas: redes sociales, compras, actividad bancaria, sistemas de control, almacenamiento en la nube y demás. Inclusive desde el año 2011, el 75% de la distribución de aplicaciones por

categorías correspondía a aplicaciones web [1].

Para intentar satisfacer la incesante demanda de aplicaciones de este tipo, los usuarios ocasionalmente se ven rodeados de software que no fue desarrollado siguiendo un modelo adecuado, que no ha sido codificado de manera correcta, posee un diseño defectuoso, entre otros tantos factores que vuelven a los sistemas poco robustos y con grandes falencias en seguridad [2,3].

En un contexto de conectividad permanente entre los usuarios, el concepto de "seguridad" se vuelve un factor fundamental [4]. Este motivo, entre otros, es el principal desencadenante de una disciplina que se encuentra en auge en distintos contextos de la actualidad, la misma es conocida como: *seguridad informática*.

La noción de seguridad en los sistemas de información, la cual se extiende a diversos entornos como software embebidos, entornos móviles, aplicaciones web, entre otros; se define como un grupo de componentes vinculados y parcialmente superpuestos, de los cuales se pueden destacar las siguientes particularidades [5]:

Seguridad: envuelve características en concordancia con la protección del sistema, sus aplicaciones y los recursos compartidos. Incluye la prevención de la adquisición y modificación no autorizada de información, es decir, se intenta cubrir tanto la seguridad del sistema, como la de los datos del usuario.

Fiabilidad, disponibilidad y recuperación: indican, respectivamente, la exactitud de la seguridad del sistema y otras funciones del mismo, el soporte a través del tiempo de un sistema seguro y la capacidad de recuperación del mismo luego de haber sufrido un ataque o algún tipo de error/accidente.

Auditabilidad: implica el seguimiento de la continua existencia de seguridad y fiabilidad del sistema, incluyendo la detección de anomalías o un comportamiento amenazante.

Un camino factible para conseguir como producto final un software que provea mecanismos de seguridad adecuados, es seguir un enfoque correctivo [6]. Esto quiere decir, llevar a cabo la evaluación de la seguridad de los sistemas informáticos a posteriori de su desarrollo. En este sentido, es preciso hacer referencia a un conjunto de herramientas denominadas: "*Herramientas de Análisis de Seguridad de Software*".

El objetivo de dichas herramientas de análisis es encontrar aquellos puntos donde el sistema informático sufre problemas relacionados con la seguridad o inclusive, identificar vulnerabilidades que podrían ser potencialmente conflictivas [7,8]. Para llevar a cabo el proceso de análisis, estas herramientas hacen uso de diferentes técnicas, por lo general su modus operandi es simular ataques en contra de la aplicación en tiempo de ejecución y, acorde a las respuestas obtenidas, generar los reportes necesarios. No obstante, se dan situaciones donde para buscar falencias específicas, la herramienta también debe ser capaz de observar errores de codificación de software [9,10].

Sobre el marco de las observaciones realizadas, es inevitable pensar que existe un amplio abanico de posibilidades a la hora de elegir una

herramienta que mejor se ajuste en una situación en particular. Por lo tanto, se vuelve preponderante contar con mecanismos capaces de evaluar las distintas alternativas accesibles.

Líneas de Investigación y Desarrollo: Análisis Multicriterio y Herramientas de Análisis de Seguridad

El área de toma de decisiones bajo criterios múltiples, o *Multiple-criteria decision-making* en su traducción al inglés (MCDM), explora un conjunto de métodos y procedimientos mediante los cuales, contextos de conflictos de esta índole, pueden ser formalmente incorporados a un proceso analítico.

Dentro de esta área de investigación, se encuentra la evaluación multicriterio. La misma propone dar solución a problemas donde se involucran un número finito de alternativas explícitamente conocidas antes de comenzar el proceso. Este tipo de análisis, ofrece un modelo para plasmar la opinión del encargado de tomar la decisión de la forma más crítica posible, es por ello que la utilidad de la alternativa elegida debe quedar representada por su desempeño en los criterios seleccionados para la evaluación [11].

De acuerdo con Huang et al. [12] en la última década, dentro de los métodos más utilizados se encuentran: MAUT (*Multi-Attribute Utility Theory*), AHP (*Analytical Hierarchy Process*) y *Outranking Methods*. Donde este último grupo incluye: PROMETHEE (*Preference Ranking Organization Method for Enrichment Evaluation*) y ELECTRE (*ELimination and Choice Expressing Reality*).

Todas estas propuestas abordan el problema de toma de decisiones de una manera similar. Esto quiere decir hacen uso de elementos matemáticos para asignar valores a las comparaciones, se ponderan "pesos" y luego se combinan estos resultados parciales para producir un resultado final.

De cualquier forma, sin importar que medio se utilice para realizar la comparación, es importante sentar las bases de específicamente que características influyen en la elección de una herramienta de análisis de seguridad.

De acuerdo con el Instituto SANS [13], en el ámbito de aplicaciones web, estas herramientas se pueden clasificar, teniendo en cuenta la forma que utilicen para llevar a cabo sus procesos, en al menos tres categorías: i) bloqueo de ataques: basados en la red de datos y en el servidor, ii) eliminación de vulnerabilidades y iii) soporte seguro para usuarios autorizados.

Esto impulsa la idea de que existe una gran variedad de posibilidades (con distintas prestaciones) a la hora de seleccionar una herramienta que ejecute un proceso de análisis de seguridad para aplicaciones web y es la principal razón de que surja la necesidad de relevar el estado del arte en cuanto a la evaluación de analizadores. En este contexto, el objetivo es conseguir un conjunto de características que pueda representar el comportamiento deseable en este tipo de software. Para concluir esta sección cabe destacar que esta línea de investigación se divide en dos temáticas relevantes: i) el estudio y selección de un método de análisis multicriterio adecuado y ii) el análisis y formación de un conjunto de criterios que sirvan para englobar todos los atributos que se deben tener en cuenta

a la hora de llevar a cabo un proceso de evaluación sobre herramientas de análisis de seguridad.

Resultados Obtenidos y Objetivos

A continuación se mencionan los resultados obtenidos hasta el momento dentro del marco de la línea de investigación planteada.

Por un lado, dentro del estudio y selección de un método de evaluación, se optó por elegir el Proceso de Análisis Jerárquico, conocido por sus siglas AHP.

Para entrar un poco en contexto, AHP hace uso de comparaciones por pares de cada elemento del problema bajo la interrogativa de: ¿Qué tantas veces es más importante un elemento por sobre otro? Y como consecuencia, el procedimiento de comparación se vuelve más simple y flexible [14]. Además, proporciona métodos para computar que tan consistentes (o no) están siendo los resultados que se van obteniendo, lo que da como producto final del proceso, tanto prioridades globales para cada una de las alternativas (posibles candidatos) involucradas, como así también medidas de consistencia para los juicios realizados.

En síntesis, la elección del método AHP por sobre otros, se justifica bajo las siguientes características: i) garantiza el tratamiento de todas las aristas involucradas en el proceso, de forma centralizada e independiente, a través del uso de una jerarquía para organizar el problema, ii) brinda herramientas para la comprobación del nivel de consistencia de los juicios realizados y por lo tanto, la calidad de los resultados obtenidos, iii) es relativamente sencillo de implementar y no precisa de muchos cálculos para obtener resultados y por último, iv) es un método muy utilizado a la hora de llevar a cabo problemas de decisión bajo criterios múltiples.

Por otro lado, en cuanto a la selección de atributos deseables, se hizo un relevamiento del estado del arte respecto a la evaluación de herramientas de análisis de seguridad, a partir del cual se pudo construir un árbol de criterios. Es importante remarcar que el árbol construido es una adaptación de los criterios expuestos en el "*Web Application Security Scanner Evaluation Criteria*" [9], el cual sirvió de base para dar un enfoque claro a la evaluación de escáneres de vulnerabilidades. Tal y como se dijo anteriormente, estas características tienen como objetivo sintetizar, de la manera más completa posible, todos los puntos que debe cubrir un escáner de seguridad para que sea lo más eficaz y eficiente posible. A continuación se expone una breve descripción de cada una:

Soporte de Protocolos: para llevar a cabo un análisis, un escáner debe admitir todos los protocolos de comunicación son comúnmente utilizados por las aplicaciones web.

Autenticación: es necesaria la compatibilidad con métodos de autenticación estándar o ampliamente utilizados para poder probar de manera efectiva las aplicaciones que requieren autenticación.

Manejo de Sesiones: durante una búsqueda de vulnerabilidades web, es importante que los escáneres mantengan sesiones válidas con la aplicación en todo momento.

Buscador: este proceso consiste en recorrer recursivamente todos los enlaces que se vayan presentando a partir de la página principal de la aplicación. Los rastreadores le deben permitir al usuario definir una gran cantidad de criterios para garantizar una exploración exhaustiva y eficiente.

Parsing: para escanear completamente una aplicación web en busca de problemas de seguridad, un escáner primero debe ser capaz de reconstruir en una imagen los aspectos arquitectónicos y funcionales de la misma.

Testing: este criterio incluye los tipos de vulnerabilidades que un escáner de aplicaciones web debería ser capaz de detectar, así como las opciones de configuración y personalización relacionadas con las pruebas que debería proporcionar.

Reportes: lo que queda luego de un proceso de análisis son las conclusiones. Para conseguir un buen resultado, es necesario que los reportes que se produzcan de la forma más clara y concisa que se pueda.

Ya definido el árbol de criterios, en relación, tanto al estudio de la toma de decisiones multicriterio, como a la seguridad informática, en el marco de esta investigación se llevó a cabo el desarrollo de "*DAST: Decision Analysis Software Tool*", una aplicación web que implementa el Proceso de Análisis Jerárquico.

Esta herramienta ofrece prestaciones que le permiten resolver el problema de toma de decisión planteado, de forma semi-automática. Por un lado, teniendo disponible para el uso un árbol de criterios (detallado en la sección anterior) para evaluar herramientas de análisis de seguridad y por otra parte, simplificando la tarea de realizar los cálculos matemáticos subyacentes al método AHP. De cualquier forma, es importante remarcar que DAST está pensada para usuarios que deseen resolver un problema de toma de decisiones bajo múltiples criterios de cualquier índole, no sólo referido a analizadores de seguridad.

En cuanto a objetivos en el corto/mediano plazo, se pretende: i) dejar a DAST funcionando como una herramienta "*as a service*" y conseguir la fácil integración con otros sistemas que lo usen para resolver problemas de toma de decisión. Esta tarea implica definir una firma y dejar el sistema subido en producción para que pueda ser accedido por el público. ii) Mejorar el Árbol de Criterios para el Dominio de Escáneres de Seguridad incluyendo otros criterios importantes (veracidad, performance, entre otros). Y por último, iii) llevar a cabo evaluaciones sobre distintas herramientas de análisis de seguridad que se utilizan actualmente en la industria.

Formación de Recursos Humanos

Las tareas llevadas a cabo en la presente línea de investigación están siendo desarrolladas como parte de un Proyecto Final Integrador en Ingeniería en Informática en la Universidad Nacional de San Luis. Se pretende que los resultados obtenidos durante el desarrollo de las tesis den origen a estudios de posgrado de los integrantes del proyecto en el que se encuentra enmarcada dicha línea de investigación. Permitiendo, de esta manera, el crecimiento académico de los integrantes de la línea como así también la generación de proyectos de investigación basados en seguridad informática.

Bibliografía

[1] VERACODE. State of software security report. The Intractable Problem of Insecure Software. 2011.

[2] PRESSMAN, Roger S. Software engineering: a practitioner's approach. Palgrave Macmillan, 2005.

[3] BLACK, Paul E. Software Assurance Metrics and Tool Evaluation. En Software Engineering Research and Practice. 2005. p. 829-835.

[4] MITRE. The standard for information security vulnerability names. Common Vulnerabilities and Exposures, 2005.

[5] SONG, Jihong; HU, Guiying; XU, QuanSheng. Operating system security and host vulnerability evaluation. En 2009 International Conference on Management and Service Science. IEEE, 2009. p. 1-4.

[6] NEUMANN, Peter G. Computer system security evaluation. En National Computer Conference. 1978.

[7] WEB APPLICATION SECURITY CONSORTIUM, et al. Web application security scanner evaluation criteria. Version, 2009, vol. 1, p. 1-26.

[8] BLACK, Paul E., et al. Software assurance tools: Web application security scanner functional specification version 1.0. Special Publication, 2008, p. 500-269.

[9] ARTHO, Cyrille; BIERE, Armin. Combined static and dynamic analysis. Electronic Notes in Theoretical Computer Science, 2005, vol. 131, p. 3-14.

[10] GANDIBLEUX, Xavier (ed.). Multiple criteria optimization: state of the art annotated bibliographic surveys. Springer Science & Business Media, 2006.

[11] HUANG, Ivy B.; KEISLER, Jeffrey; LINKOV, Igor. Multi-criteria decision analysis in environmental sciences: Ten years of applications and trends. Science of the total environment, 2011, vol. 409, no 19, p. 3578-3594.

[12] VAIDYA, Omkarprasad S.; KUMAR, Sushil. Analytic hierarchy process: An overview of applications. European Journal of operational research, 2006, vol. 169, no 1, p. 1-29.

[13] HARDIKAR, A.; BAMBENEK, J. C. A. Malware 101—Viruses. SANS Institute InfoSec Reading Room, 2008.

Experiencia de Usuario en Plataforma virtual de Aprendizaje

Iván Balmaceda Castro[1], Carlos Salgado[2], Mario Peralta[2], Alberto Sánchez[2], Mariela Fernandez[1], Juan Magaquian[1], Nelson Fuentes[1]
[1]Departamento Académico de Ciencias Exactas, Físicas y Naturales
Sede Regional Chamical. Universidad Nacional de La Rioja
e-mail: {ibalmaceda89, mariel.lucia, jmagaquian, neluis97}@gmail.com,
[2]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
{csalgado, mperalta, alfanego}@unsl.edu.ar

RESUMEN

La Experiencia de Usuario puede considerarse como el proceso que lleva a cabo el usuario cuando interactúa con un producto. Las plataformas virtuales de aprendizaje, permiten que se articulen e integren diferentes métodos y estrategias de enseñanza. Así, existen desafíos relacionados al desarrollo y aprovechamiento de dichas plataformas que consisten en aprovechar la información proporcionada por los indicadores de uso para monitorizar y analizar la interacción en el proceso de enseñanza y aprendizaje para detectar dificultades y deficiencias en el esquema de trabajo. Como en cualquier entorno virtual, el diseño de estas plataformas, también deben centrarse en el usuario, ser usable y considerar las características y habilidades de los usuarios a la hora de interaccionar con la interfaz gráfica y los contenidos educativos.
Así, la UX no puede reducirse al estudio de lo que hace el usuario, además es necesario comprender por qué lo hace. En este sentido, se está trabajando en la definición de métricas de usabilidad y accesibilidad que permitan generar una Heurística que sirva de instrumento de Evaluación de la Experiencia de Usuario en plataformas virtuales de aprendizaje, a partir de la identificación de algunos criterios que permiten cuantificar el grado de satisfacción de los usuarios.

Palabras clave: UX, Experiencia de Usuario, Plataformas Virtuales de Aprendizaje, Métricas.

CONTEXTO

El presente trabajo se enmarca dentro del proyecto de investigación "Modelo de Proceso y Evaluación Centrado en el Usuario incorporando requisitos de Usabilidad y Experiencia de Usuarios" - Centro de Investigación y Desarrollo Informático, Departamento Académico de Ciencias Exactas, Físicas y Naturales - Sede Regional Chamical Universidad Nacional de La Rioja. Cod. Proy. 27PIN/C0003, financiado por Consejo de Investigaciones Científicas y Tecnológicas UNLaR Res N°052/18. Y con la colaboración 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 Nº P-031516. Dicho proyecto se encuentra reconocido por el programa de Incentivos.

1. INTRODUCCIÓN

El desarrollo de sistemas de información 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 (IS) es construir aplicaciones de calidad, útiles a los usuarios finales, aplicando distintos métodos y principios. El diseño centrado en el usuario surge como un enfoque y método que consiste en conocer algunas particularidades del usuario, con el objetivo de hacer más familiares y efectivas las interfaces graficas que se diseñan para él. Esto permite encontrar una solución visual y

funcional a diversos requerimientos comunicativos y funcionales en sistemas informáticos. El estándar internacional ISO 13407, describe cómo un proceso de Diseño Centrado en Usuario (DCU) puede ser usado para conseguir sistemas usables. Este, provee un marco de trabajo para aplicar técnicas en el diseño y evaluación del DCU. La Usability Professionals' Association (UPA) lo define como un enfoque de diseño cuyo proceso está dirigido por información sobre las personas que van a hacer uso del producto. Los procesos DCU se centran en los usuarios a través de la planificación, el diseño y el desarrollo de un producto, [1, 2, 3].

Los productos con plataforma en la Web constituyen una herramienta muy poderosa a la hora de acercar propuestas de información, comunicación, productos y servicios a los ciudadanos. Como se mencionó previamente, la Experiencia de Usuario (UX por sus términos en inglés User eXperience) puede considerarse como el proceso que lleva a cabo el usuario cuando interactúa con un producto. A veces se confunden estos conceptos con usabilidad, ésta es la facilidad y satisfacción con la que pueden utilizar un producto, mientras que en la experiencia de usuario se involucran factores sociales, culturales y de contexto que pueden generar una percepción positiva o negativa del producto [4, 5, 6, 7, 8]. UX es entendido como dinámico, da el estado emocional y cambiante de una persona y las diferencias en las circunstancias durante y después de una interacción con un producto, buscando resolver el problema de la utilidad del producto en la satisfacción de su uso. [7, 9, 10, 11]

Las plataformas virtuales de aprendizaje, permiten que se articulen e integren diferentes métodos y estrategias de enseñanza, tales como el aprendizaje colaborativo y la interacción de los participantes. En la actualidad existen varios desafíos relacionados estrechamente con el desarrollo y aprovechamiento de las plataformas virtuales de aprendizaje que requieren una especial atención. Estos consisten en aprovechar la información proporcionada por los indicadores de uso para monitorizar y analizar la interacción en el proceso de enseñanza y aprendizaje con el fin de detectar dificultades y deficiencias en el esquema de trabajo. Esta interacción está enfocada principalmente a mejorar la experiencia de uso y resultados académicos del alumnado [12, 13, 14]. A su vez, se centra en las habilidades y procesos cognitivos del usuario, estudiando únicamente su comportamiento racional y dejando de lado su comportamiento emocional [15, 16].

Tal como ocurre en cualquier entorno virtual, el diseño de las plataformas virtuales de aprendizaje, también deben ser centrado en el usuario [17], usable y que tenga en cuenta las características y habilidades de los usuarios a la hora de interaccionar con la interfaz gráfica y con los contenidos educativos.

Es así que, la UX no puede reducirse únicamente al estudio técnico de lo que realiza el usuario. Por consiguiente, su medición no debe quedar en una mera descripción de lo que hace. Es necesario comprender por qué lo hace para determinar su experiencia. Para ello, es necesario saber cómo se sienten las personas cuando usan un sistema, considerando la usabilidad, la satisfacción, entre otros. La motivación y expectativas del usuario tienen un papel muy importante en la experiencia del usuario en lugar de la usabilidad tradicional. [11, 18, 19].

A través de estudios preliminares, se analizaron con profundidad atributos de usabilidad y técnicas aplicables a lo largo del ciclo de vida de desarrollo del software [20], que permitió integrar la Ingeniería de Usabilidad con la Ingeniería de Software, a través de metodologías que lo propicien.

También se analizaron métodos, técnicas y herramientas para cada etapa del DCU [21], que involucró a personas con discapacidad visual, las que permitieron percibir, entender, navegar e interactuar con la web, aportando a su vez contenidos.

Los métodos de inspección basados en heuristicas [22], han pemitido evaluar la forma en que el sistema interactúa con el usuario a fin de determiner cómo se realizan las tareas básicas.

En este sentido, se está trabajando en la

definición de métricas de usabilidad y accesibilidad que permitan generar una Heurística que sirva de instrumento de Evaluación de la Experiencia de Usuario en plataformas virtuales de aprendizaje, a partir de la identificación de algunos criterios que permiten cuantificar el grado de satisfacción de los usuarios de este tipo de herramientas.

Ésta, medirá la calidad de la interfaz en relación a su facilidad para ser aprendido y usado por usuarios en un determinado contexto de uso.

Actualmente en la UNLAR se utilizan dos plataformas: 1) EVA UNLaR, que se utiliza como complemento a las carreras presenciales 2) Virtual UNLaR para las carreras de Cs de la Educación que son semi presenciales. Dichas plataformas, serán utilizadas para la validación de la propuesta, a fin de verificar el grado de usabilidad, identificar mejoras y, por sobre todo, mejorar la experiencia del usuario en las mismas para que la capacitación sea de mayor calidad educativa.

Para ello se tomarán como guía el listado de principios Heurísticos de Jakob Nielsen [18], estas son generales y se pueden aplicar a cualquier sitio web.

Se puede afirmar que cuando se trata de mejorar la UX en la Web, no sólo la Usabilidad es un componente clave, también lo es la Accesibilidad web, otro factor de calidad [4].

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En esta línea de investigación se trabaja en lo referente a distintos modelos y métodos de evaluación de calidad en la experiencia de usuario. Así, acorde con los objetivos planteados en el proyecto "Modelo de Proceso y Evaluación Centrado en el Usuario incorporando requisitos de Usabilidad y Experiencia de Usuarios" perteneciente a la Universidad Nacional de La Rioja, sede Chamical, La Rioja, los ejes principales de la investigación se centran en el análisis de las pautas de Ingeniería de Usabilidad, análisis y definición de requerimientos funcionales y no funcionales basados en dichas pautas, definición e integración de un modelo que incluya los requisitos de experiencia de usuario y usabilidad, como la definición de Métricas de Usabilidad que permitan evaluar dicho modelo y su validación a través de un caso de estudio.

3. RESULTADOS OBTENIDOS Y OBJETIVOS

El objetivo general de la presente línea de investigación consiste en incorporar requisitos de usabilidad y experiencia de usuario en un modelo de proceso y evaluación centrado en el usuario para mejorar la usabilidad de un software. En este contexto, se plantean los siguientes objetivos particulares:

1. Analizar las pautas de Ingeniería de Usabilidad.
2. Analizar y definir los requerimientos funcionales y no funcionales basada en las pautas de Usabilidad.
3. Analizar en detalle las metodologías de desarrollo de software, específicamente en los modelos, tipos, estrategias de adquisición de requisitos de software.
4. Integrar y Definir un modelo que incluya los requisitos de experiencia de usuario y usabilidad.
5. Desarrollar Métricas de Usabilidad que permitan evaluar el modelo.
6. Validar el modelo en un caso de estudio.

El diseño de interfaces de usuario es una tarea que ha adquirido relevancia en el desarrollo de un sistema. La calidad de la interfaz de usuario puede ser uno de los motivos que conduzca a un sistema al éxito o al fracaso.

En este sentido se está trabajando en la definición de un modelo de calidad pensando en las experiencias de los usuarios con respecto a productos software. Para ello, se tiene en cuenta, además de las guías de experiencias de usuario, las características sugeridas por la Norma ISO 25010. Los principios que se presentan son de utilidad para creación de interfaces funcionales y de fácil operación.

Muchas veces es necesaria la combinación de diversos instrumentos, como la prototipación y la aplicación de heurísticas de evaluación para facilitar el proceso de diseño.

Por ello, la definición de un modelo de calidad, basado en un conjunto de características y subcaracterísticas que deberían satisfacer las interfaces de usuario como, por ejemplo:

Claridad: significa que la información es transmitida de manera precisa.

Concisión: Si bien la claridad es importante, no significa que sea necesario agregar información para asegurarse que ésta llegue al usuario de manera correcta. Cualquier otra información que el usuario necesite, puede revisar documentación adicional, pero la información debe ser clara y concisa.

Coherencia: Ayuda a los usuarios a desarrollar patrones de uso, y a aprender la función de los diversos elementos propios de la interfaz. El usuario debe sentirse cómodo con dicha interfaz.

Legibilidad: El lenguaje utilizado debe ser simple para ayudar a la rápida lectura por parte del usuario. Esto ayudará también a que el usuario se acostumbre más rápidamente a las diversas funciones que posee la interfaz.

Interactividad: Una buena interfaz tiene que ser rápida. Además, la interfaz debería contextualizar al usuario respecto de las tareas que se están realizando.

Flexibilidad: La interfaz de usuario debería permitirle deshacer las acciones erróneas que el usuario pudiera realizar. Por ejemplo, posibilitar al usuario la restauración de elementos que hayan sido borrados.

Eficiencia: La interfaz de usuario debería entender el objetivo del usuario y no impedírselo de ningún modo con, por ejemplo, opciones adicionales sobre lo que desea hacer.

Estética: Dado que el diseño de la interfaz se centra en el usuario, se tiene que tener en cuenta el atractivo visual. Cada vez más usuarios están pendientes del diseño, y no solo buscan funcionalidad sino también algo visualmente atractivo.

El diseño de una interfaz intuitiva, fácil de usar, rápida y de diseño atractivo, aumenta la probabilidad de que los usuarios la utilicen y, además, de lograr una fidelización con la aplicación software.

4. FORMACIÓN DE RECURSOS HUMANOS

Se está trabajando en el Proyecto de Tesis de la Maestría en Ingeniería de Software (Plan Ord. 005/10-CD) del Lic. Iván Balmaceda Castro, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Sistemas de Información.

Becarios de Grado: a designar en la convocatoria a becas Ayudante de Investigación 2019 de la Universidad Nacional de La Rioja.

5. REFERENCIAS

1. Constantine, L.L., Loockwood L.A.D., “Software for Use: A practical guide to the Models and Methods of Usage-Centered Design”. Addison-Wesley (1999).
2. Hassan, Y; Ortega Santamaria, S. (2008) Informe APEI sobre Usabilidad. APEI (versión online: http://www.nosolousabilidad.com/manual/3.htm).
3. UPA. Usability Professionals’ Association. What is User-Centerd Design? (versión online http://www.upassoc.org/usability_resources/about_usability/what_is_ucdhtml).
4. Martin, G. Gaetaán, V. Saldaño, A. Pires, G. Miranda, S. Villagra, A. Carrizo, C. Cardozo, H. Sosa “Un enfoque integrador para diseñar y evaluar interfaces de usuario web” XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires), 2007

5. ISO 9241-11, Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts, International Organization for Standardization, 2018.
6. J. Nielsen, “Usability inspection methods,” in Conf. Human factors in computing systems, Boston, Massachusetts, United States, 1994.
7. J. Nielsen, and R. Molich, “Heuristic Evaluation of User Interfaces”, in Proceeding SIGCHI '90 Conference on Human factors in Computing Systems, pp.249-256, 1990.
8. N. Bevan, “International Standards for HCI and Usability”, International journal of human-computer studies, vol. 55, no. 4, pp. 533-552, 2001.
9. H. Hartson, and T. Andre, “Criteria for Evaluating Usability Evaluation Methods” In International Journal of Human-Computer Interaction, pp. 373-410, 2003.
10. ISO/IEC 9126-1, Software Engineering - Product Quality, 2001.
11. E. Sosa Bruchmann, G. Montejano, A. Garis “Análisis de la experiencia del usuario: relación entre el comportamiento emocional y la satisfacción de uso” XVII Workshop de Investigadores en Ciencias de la Computación, 2015.
12. A. Allam, A. Razak and H. Mohamed, "User Experience: Challenges and Opportunities", Journal of Information Systems Research and Innovation, vol. 3,pp. 28-36, 2013.
13. A. Camacho-Navarro “El docente como agente implicado en crear Comunidades Virtuales de Aprendizaje”. Revista electrónica en Ciencias Sociales y Humanidades Apoyadas por Tecnologías, 7(2), 57-68. 2019.
14. M. Cantabella, B. López-Ayuso, A. Muñoz and A. Caballero “Una herramienta para el seguimiento del profesorado universitario en Entornos Virtuales de Aprendizaje.” Revista Española de Documentación Científica, 2016.
15. Djajadiningrat, J.P.; Overbeeke, C.J.;Wensveen, S.A.G. (2000). “Augmenting Fun and Beauty: A Pamphlet”. DARE 2000, Elsinore, Denmark. Disponible en: http://www.itproducts.sdu.dk/djajadiningrat/publications/00DjajDARAugm.pdf
16. Alva, M. (2005). “Metodología de Medición y Evaluación de la Usabilidad en Sitios web Educativos”. Tesis Doctoral, Universidad de Oviedo
17. J. T. Hackos, J. Redish, User Interface Task Analysis, John Wiley & Sons, 1998.
18. J. Nielsen, “Usability engineering”, 1 edition ed., San Diego, CA 92101, UnitedStates: Academic Press, 1993.
19. Carreras, J. (2003). “Midiendo la experiencia de usuario”.
20. M. Mascheroni,C. Greiner, R. Petris, G. Dapozo, M. Estayno “Calidad de Software e Ingenieria de Usabilidad” Workshop de Investigadores en Ciencias de la Computación (WICC 2012, UNM, Misiones), 2012
21. G. Lafuente, C. Ballesteros, J. Filippi “Desarrollo de una Plataforma de Capacitación a Distancia Utilizable y Abierta para Personas con Discapacidad Visual” Workshop de Investigadores en Ciencias de la Computación (WICC 2018, UNN, Corrientes), 2018
22. A. Ferreira, C. Saenz “Usabilidad de los entornos virtuales de enseñanza y aprendizaje” VI Congreso de Tecnología en Educación y Educación en Tecnología, 2011.

Fortalecimiento de la calidad en procesos de software y Gobernanza Digital

Silvia Esponda, Ariel Pasini, Marcos Boracchia,
Julieta Calabrese[1], Rocío Muñoz[1], Juan Santiago Preisegger[1], Patricia Pesado

Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
[1] Becario postgrado UNLP
50 y 120 La Plata Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 La Plata Buenos Aires
(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 de tópicos concernientes con la mejora de los procesos de gestión, la calidad en los procesos de desarrollo y productos de software, y e-goverment.
En este ámbito, se han investigado y desarrollado distintos servicios de gobierno digitales (con foco en políticas de prestación de servicios y herramientas de uso que beneficien a la comunidad), se ha trabajado en la mejora de procesos de 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, haciendo hincapié en los datos utilizados.

Palabras Claves
Ingeniería de Software – Calidad del proceso – Calidad del producto - Normas de Calidad – ISO - Gobernanza Digital – Blockchain

CONTEXTO

Esta línea de investigación se enmarca en el proyecto "11/F023 Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso" y en el subproyecto "Gobernanza Digital. Mejora de Procesos" (2018-2021), acreditado por el Ministerio de educación de la Nación.
Además, la línea se encuentra abordada por el proyecto "Calidad de Datos y Blockchain" aprobado por la Facultad de Informática UNLP.
El III-LIDI participa en el Proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de Buenos Aires (PITAP-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 del Sur (UNS).
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 y participa en proyectos internacionales, entre ellos ERASMUS "Fortalecimiento de la capacidad de gobernanza de ciudades sostenibles inteligentes".

Desde el año 2015 el Instituto es Centro Asociado de la CIC.

1. INTRODUCCION

El uso de diferentes Tecnologías de Información y Comunicación (TICs) ha modificado el desarrollo de un gran número de actividades cotidianas.
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. Dado que se toman muchas decisiones basadas en dicha información, se debería asegurar que es confiable.
Una ciudad digital es aquella en la que, utilizando los recursos propios de la infraestructura de telecomunicaciones y de la informática existentes, brinda a sus habitantes un conjunto de servicios digitales a fin de mejorar el nivel de desarrollo humano, económico y cultural de esa comunidad a nivel individual y 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 por una organización (con un cierto control) destinada a satisfacer necesidades de la comunidad." 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. Existen diferentes estrategias de enfocar la calidad de estos servicios y 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: Dentro de este grupo se cuenta con una norma específica para evaluar la calidad de los datos utlizados, 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 el proyecto PITAP-BA "Herramientas para el desarrollo y la entrega de servicios públicos digitales de acción social para municipios bonaerenses", en conjunto con el LISSI-UNS. Dicho proyecto busca mejorar la eficacia de la implementación de las políticas sociales municipales mediante soluciones informáticas, con el fin de mejorar la entrega de servicios públicos de acción social.

En el marco de este proyecto, el III-LIDI se enfocó en la investigación sobre las características de gobierno abierto y su aplicación en las distintas agencias gubernamentales, como así también en el diseño de soluciones y estrategias para integrar diferentes actividades mejorando la eficacia y eficiencia de las interacciones gobierno-ciudadanos.
Como resultado se definió un modelo de evaluación de gobierno abierto, que se aplicó a los municipios de la provincia de Buenos Aires y puede expandirse a las capitales provinciales de nuestro país. Con respecto al diseño de soluciones se presentó una herramienta para la generación de portales web que permitan

brindar, a un conjunto de receptores, la cartera de servicios que un organismo pone a su disposición. Dicha herramienta brinda la posibilidad de establecer un formato a ser respetado, teniendo en cuenta tanto la presencia institucional como el contenido de los servicios.

El III-LIDI trabaja, desde el año 2003 en aplicaciones relacionadas con la gobernanza digital para la elección de autoridades mediante el voto electrónico. Se destacan la definición e implementación de tres modelos (presencial, semipresencial y remoto) en 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 Gestión Electrónica Digital y 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". En 2018 se realizó la migración de la certificación a la versión IRAM-ISO 9001:2015 y se actualizo el alcance del SGC.
Asimismo en 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 2018 se realizó la migración de la certificación a la versión IRAM-ISO 9001:2015.

3- Mejora de los servicios de gobierno digital en unidades académicas de nivel universitario

En el marco de las mejoras 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. [2]

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 de 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”. [3]

La Facultad de Informática posee un Portal de Gestión Administrativa que provee a sus Alumnos, Docentes, No Docentes y Graduados una guía para la realización de trámites de los principales servicios que presta. De cada uno de estos trámites se brinda información estandarizada sobre: Descripción, Forma de Acceso, Requisitos y Contacto.

4- Mejora de Procesos de gestión en el desarrollo de software

La gestión de la calidad es una actividad fundamental dentro de cualquier organización. Dicha gestión puede llegar a ser una ventaja competitiva que la fortalezca en el momento de prestar un mejor servicio o de tener un producto que cumpla las exigencias y expectativas del cliente. En el camino de obtener una buena gestión de la calidad de procesos en las PyMEs, es necesario establecer un Sistema de Gestión de la Calidad (SGC) de acuerdo a la norma IRAM-ISO 9001.

Bajo este contexto, la norma IRAM-ISO 9001:2015 "Sistemas de gestión de la calidad. Requisitos" [4] se funda en los retos que enfrentan las empresas de cualquier tamaño y sector hoy en día, centrándose en la eficacia del sistema de gestión para dar cumplimiento a los requisitos del cliente. La norma promueve la adopción de un enfoque basado en procesos y exige cierta información documentada que la organización determine como necesaria para la eficacia del SGC.

Por otro lado, las metodologías ágiles representan una alternativa muy utilizada en la actualidad para el proceso de desarrollo de sistemas de software que hace hincapié en la relación con el cliente y en el desarrollo incremental del producto. Una de las metodologías más utilizadas en la actualidad, Scrum, se define como un proceso iterativo incremental y empírico para administrar y controlar el trabajo de desarrollo.

En el camino de obtener una buena gestión de la calidad de procesos en las PyMEs que utilizan Scrum en sus procesos de desarrollo, surgió la necesidad de compatibilizar la documentación utilizada por la metodología con la documentación requerida por la norma para la eficacia del SGC. El resultado de un análisis de esta situación concluyó que la documentación definida por la metodología no era suficiente para satisfacer los requisitos del estándar IRAM-ISO 9001:2015.

En base a esta problemática, se generó un documento único integrador que reúne un conjunto de ítems basado en estándares IEEE adicionales a la documentación de Scrum. La integración buscó satisfacer la documentación necesaria definida en los requisitos del estándar IRAM-ISO 9001:2015 (en las secciones planificación, soporte y operación) a la hora de crear un producto de software utilizando la metodología ágil de desarrollo Scrum. El estudio de estándares IEEE estuvo enfocado en el documento de Especificación de Requisitos de Software (SRS) y en el documento de Plan de Gestión de Proyectos (PGP).

5- Calidad en productos - Calidad de datos.

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. [5]

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. Uno de los pasos de ésta, consiste en seleccionar las características a evaluar, las cuales se encuentran definidas en el

modelo de calidad propuesto por la norma ISO/IEC 25010.[6], [7]

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), que si bien no asegura una certificación, puede ser utilizada como guía por personas no experimentadas.

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.[8][9]

Por otro lado, con la aparición de la tecnología Blockchain es de interés estudiar los temas de calidad asociados.

Blockchain surge para eliminar intermediarios o entidades centralizadas para la realización y almacenamiento de dichas transacciones, reduciendo así los riesgos de errores humanos.

Esta tecnología consta de tres componentes fundamentales: una transacción, un registro de transacciones (formando un bloque) y un sistema que verifica y almacena todos los bloques en la cadena. Una característica fundamental de esta tecnología es que se compone y mantiene mediante una red distribuida de "nodos" que tienen acceso a participar en la validación de bloques y transacciones, como así también a almacenar nuevas transacciones en los bloques de la cadena. Cada transacción está compuesta por una serie de datos que representan la información específica de la misma.

Es fundamental contar con datos que sean correctos, completos y consistentes. Existen normas de calidad, entre ellas la ISO/IEC 25012, para evaluar la calidad de los datos. Las distintas características que presentan y proponen las normas podrían no alcanzar para asegurar que se cumplan los objetivos perseguidos por esta nueva tecnología.

Se propone analizar e investigar las especificaciones que contienen los datos manejados por dicha tecnología y evaluar la aplicación de dicha norma.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Análisis y estudio de normas y modelos orientados a la Calidad del Producto, como por ejemplo la familia IRAM - ISO/IEC 25000
- Evaluación de procesos en organismos públicos y privados según los requisitos de IRAM - ISO 9001.
- Evaluación de procesos de software según las directrices de la ISO/IEC 90003
-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.

3. RESULTADOS OBTENIDOS/ ESPERADOS

- Se avanza en la validación del modelo de evaluación de servicios universitarios con la información relevada de las diferentes unidades académicas

- Se avanza en la extensión del modelo de evaluación de gobierno abierto, aplicado a la provincia de Buenos Aires y a las capitales del País
- Se desarrolló una herramienta para la generación de portales WEB que presenten los servicios de una institución gubernamental.
- Se continua con el mantenimiento de los procesos certificados bajo la norma IRAM-ISO 9001:2015: "Pre-Ingreso a Distancia de la Facultad de Informática" y "Concursos Docentes de la Facultad de Informática".
- Se analizan 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.
- Se generó un Portal de Gestión Administrativa para la Facultad de Informática UNLP, donde se reúne información y enlaces sobre los diferentes servicios que brinda la facultad.
- Se propuso un documento único que integra documentación utilizada en estándares conocidos y en la metodología ágil Scrum, con el fin de satisfacer la información documentada solicitada por la norma IRAM-ISO 9001:2015 para la eficacia del SGC.
- Capacitación y desarrollo de los documentos básicos de gestión de la calidad de productos.
- Se realizaron acciones de consultoría 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".
- Se aplicó el Voto electrónico presencial en elecciones universitarias y Voto por Internet en distintos organismos.

4. FORMACION DE RECURSOS HUMANOS

- Capacitación de los miembros del proyecto a través de diversos cursos del Instituto Argentino de Normalización y Certificación (IRAM).
- El proyecto cuenta con becarios de Maestría de la UNLP en los temas del área.
- 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"

5. BIBLIOGRAFIA

[1] E. Estévez and T. Janowski, "Gobierno Digital, Ciudadanos y Ciudades Inteligentes.," 2016.

[2] 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.

[3] 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.

[4] IRAM-ISO 9001:2015 "Sistemas de gestión de la calidad. Requisitos".

[5] ISO, "ISO/IEC 25000:2014 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Guide to SQuaREtle," 2014.

[6] ISO, "ISO/IEC 25040:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- Evaluation process," 2011.

[7] ISO, "ISO/IEC 25010:2011 Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models," 2011.

[8] D. St-Louis and W. Suryn, "Enhancing ISO/IEC 25021 quality measure elements for wider application within ISO 25000 series," 2012, pp. 3120–3125.

[9] 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.

Fortalecimiento de las Capacidades de Gobernanza para Ciudades Inteligentes Sostenibles

Armando De Giusti [1], Patricia Pesado [1], Ariel Pasini [1], Pablo Thomas [1], Rocío Muñoz[1,2], Juan Santiago Preisegger[1,2], Karina Cenci[3], Carlos Chesñevar[3], Elsa Estevez[3], Pablo Fillottrani[3], Sergio Gómez[3], Sonia Rueda[3], Axel Soto[3]

[1]**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
526 e/ 10 y 11 - La Plata, Buenos Aires
[2]**Becario postgrado UNLP**
(degiusti, ppesado, apasini, pthomas, rmunoz, jspreisegger) @lidi.info.unlp.edu.ar

[3]**Laboratorio de Ingeniería de Software y Sistemas de Información (LISSI)**
Departamento de Ciencias e Ingeniería de la Computación – Universidad Nacional del Sur
Av. San Andrés 800 – Campus de Palihue - Bahía Blanca, Buenos Aires
Centro Asociado CIC
526 e/ 10 y 11 - La Plata, Buenos Aires
(kmc, cic, ece, prf, sag, svr, axel.soto) @cs.uns.edu.ar

RESUMEN

Este trabajo presenta un proyecto de investigación y desarrollo que tiene como objetivo el fortalecimiento de las capacidades de gobernanza para ciudades inteligentes en América Latina y el Caribe (ALC) a través del desarrollo de cursos y programas de postgrado. El Proyecto es co-financiado en el marco del Programa Erasmus+ de la Unión Europea, para la construcción de capacidades en la educación superior. En particular, el artículo describe los objetivos de investigación, actividades a desarrollar y entregables a producir.

Palabras Claves
Ciudades Inteligentes - Capacidades de Gobernanza - Gobernanza Digital - Colaboración Internacional

CONTEXTO

La línea de investigación y desarrollo aquí descripta es parte de las tareas planificadas por el proyecto "Strengthening Governance Capacity for Smart Sustainable Cities (CAP4CITY)". El Proyecto, financiado como parte del programa Erasmus+ de la Unión Europea es ejecutado por un consorcio integrado por 12 universidades; cuatro de ellas europeas – Donau Universität für Weiterbildung (DUK) en Austria, Tallinn University of Technology (TUT) en Estonia, Delft University of Technology (TU Delft) en los Países Bajos, y Gdańsk University of Technology (GUT) en Polonia; y ocho universidades en la región de ALC – Universidad Nacional de La Plata (UNLP) y Universidad Nacional del Sur (UNS) en Argentina; União Brasileira de Educação e Assistência (PUCRS) y Faculdade Meridional (IMED) en Brazil; Universidad Técnica Federico Santa María (UTFSM) y Universidad Católica del Norte (UCN) en Chile; y Universidad Externado de Colombia (UEC) y Escuela Colombiana de Ingeniería (ECI) en Colombia. El Proyecto lleva el número 598273 y se ejecuta bajo el acuerdo 598273-EPP-1-2018-1-AT-EPPKA2-CBHE-JP.

1. INTRODUCCION

El fin de este proyecto es utilizar la gran atención que el concepto de Ciudades Inteligentes Sostenibles (SSC) ha alcanzado en América Latina y el Caribe, así como en otras regiones [1, 2, 3, 4, 5. 6, 7, 8, 9], e integrarlo en diversos cursos universitarios utilizando nuevas e innovadoras metodologías, herramientas de enseñanza y de aprendizaje, desarrollando nuevos planes de estudio en todos los niveles del proceso educativo. El objetivo final es fortalecer las capacidades de gobernanza para facilitar la implementación de iniciativas de ciudades inteligentes.

Dado el creciente número de competencias necesarias y su característica interdisciplinaria, los planes de estudio de SSC se implementarán en diferentes áreas, tales como: Tecnologías de la Información y la Comunicación (TICs), Administración de Empresas, Informática, Ingeniería, Arquitectura y Urbanismo, Planificación Urbana, y Ciencias Políticas, entre otras. Esto se hará a través de una red colaborativa e internacional de instituciones académicas seleccionadas en América Latina y Europa, que apoya la modernización e internacionalización del campo de la educación superior en los países socios.

Se espera que los resultados del proyecto sean utilizados por:

- Estudiantes de Estudios Urbanos, Sistemas de Información, Administración de Empresas, Administración Pública, Políticas Públicas, Ciencias Políticas, Arquitectura y otras disciplinas interesadas en conocer sobre Ciudades Inteligentes Sostenibles (SSC).
- Facultades y otras unidades académicas universitarias interesadas en desarrollar y ofrecer cursos sobre SSC.
- Interesados en realizar investigaciones sobre SSC utilizando sus propios métodos y enfoques disciplinarios.
- Responsables políticos, administradores públicos, ejecutivos gubernamentales, asesores legales y otros miembros de las administraciones municipales cuyas responsabilidades involucran la planificación y gestión de iniciativas de SSC.
- Profesionales, gerentes y empresarios que trabajan para organizaciones sin fines de lucro o del sector privado que implementan iniciativas de SSC en colaboración con las ciudades.

Dentro de los receptores también se incluirán instituciones de educación superior a nivel local, provincial y nacional. Para las instituciones de educación superior y las administraciones ciudadanas a nivel europeo, un mecanismo de participación particular sería pertenecer a una red de SSC para compartir conocimientos e intercambiar experiencias sobre el desarrollo y la entrega de programas educativos relacionados con SSC.

El resto de este artículo se organiza de la siguiente manera. La Sección 2 explica los objetivos de investigación, la Sección 3 los resultados esperados y, por último, la Sección 4 discute la formación de recursos humanos.

2. OBJETIVOS DE INVESTIGACIÓN Y ACTIVIDADES

Los principales objetivos del proyecto son fortalecer y desarrollar la capacidad de instituciones académicas determinadas en América Latina y Europa para:

- mejorar la calidad de la educación superior en el campo de SSC, aumentando su relevancia para el

mercado laboral y la sociedad en general;
- aumentar las competencias en los países en desarrollo de América Latina para enfrentar los desafíos del mundo digital;
- proveer el desarrollo continuo de competencias en SSC a través de programas de capacitación y educación;
- asegurar una enseñanza multidisciplinaria orientada a problemas en SSC;
- desarrollar una red de cooperación entre los socios de diferentes regiones del mundo;
- promover la colaboración entre los socios del Consorcio, las entidades públicas, las empresas y otros interesados en SSC;
- facilitar el intercambio de conocimientos, experiencias y buenas prácticas e iniciativas conjuntas en materia de SSC entre socios académicos ubicados en diferentes ciudades y países.

A fin de contribuir con los objetivos previamente definidos, desde la Universidad Nacional de La Plata y la Universidad Nacional del Sur se colaborará con las siguientes actividades:

- relevamiento de iniciativas de SSC que puedan identificarse como buenas prácticas y de programas de postgrado relacionados con temas de SSC;
- capacitación de formadores en SSC;
- implementación de nuevos programas educacionales en SSC;
- difusión y explotación del concepto de SSC en Argentina;
- fortalecimiento de los vínculos laborales con sectores públicos y privados de la región.

3. RESULTADOS ESPERADOS

Como resultado de las tareas a desarrollar se esperan obtener los siguientes resultados:

- desarrollo de los módulos de formación relacionados con SSC y entrega online de los mismos;
- organización de dos tipos de talleres en América Latina para: 1) identificar el tipo de competencias que se necesita desarrollar y 2) validar de los módulos de formación propuestos;
- implementación de la “Capacitación de formadores”;
- relevamiento del estado actual de los programas educacionales y casos de estudio relacionados con SSC en Argentina;
- implementación de nuevos planes de estudio de postgrado relacionados con SSC en conjunto con otras universidades socias;
- incremento de la matrícula de alumnos interesados en los postgrados de SSC;
- incremento de publicaciones relacionadas con SSC por parte de las universidades involucradas;
- generación de nuevas oportunidades laborales en empresas de la región.

4. FORMACION DE RECURSOS HUMANOS

Adicionalmente a los resultados científicos y académicos, el proyecto prevee la formación de recursos humanos. Estos esfuerzos incluyen:

- capacitación de los miembros del proyecto en SSC;
- desarrollo de tesis de postgrado y tesinas de grado en el área;
- participación de los integrantes de esta línea de investigación en el dictado de asignaturas/cursos de grado/postgrado

en la Facultad de Informática de la UNLP y en el Departamento de Ciencias e Ingeniería de la Computación de la UNS.

5. BIBLIOGRAFIA

[1] Akande, A., Cabral, P., Gomes, P., Casteleyn, S. 2019. "The Lisbon ranking for smart sustainable cities in Europe." Sustainable Cities and Society, 44, pp. 475-487.

[2] Caragliu, Andrea, Chiara Del Bo, and Peter Nijkamp. 2011. "Smart Cities in Europe." Journal of Urban Technology 18(2):65–82.

[3] Craglia, Massimo, Lila Leontidou, Giampaolo Nuvolati, and Jürgen Schweikart. 2004. "Towards the Development of Quality of Life Indicators in the 'Digital' City." Environment and Planning B: Planning and Design 31(1):51–64.

[4] Elsa Estevez, Nuno Lopes, Tomasz Janowski. 2016. "Smart sustainable cities: Reconnaissance Study", 1-330.

[5] Huovila, A., Bosch, P., Airaksinen, M. 2019. "Comparative analysis of standardized indicators for Smart sustainable cities: What indicators and standards to use and when?" Cities, 89, pp. 141-153.

[6] Keshvardoost, S., Renukappa, S., Suresh, S. 2019. "Developments of policies related to smart cities: A critical review". Proceedings - 11th IEEE/ACM International Conference on Utility and Cloud Computing Companion, UCC Companion 2018, art. no. 8605807, pp. 365-369.

[7] Kurebayashi, Toshihiko, Yoshihiro Masuyama, Kiyonori Morita, Naoyuki Taniguchi, and Fumio Mizuki. 2011. "Global Initiatives for Smart Urban Development." Hitachi Review 60(2):89–93.

[8] Martin, C., Evans, J., Karvonen, A., Paskaleva, K., Yang, D., Linjordet, T. 2019. "Smart-sustainability: A new urban fix?" Sustainable Cities and Society, 45, pp. 640-648.

[9] Paroutis, Sotirios, Mark Bennett, and Loizos Heracleous. 2014. "A Strategic View on Smart City Technology: The Case of IBM Smarter Cities During a Recession." Technological Forecasting and Social Change 89:262– 72.

Guía para la aplicación de la Norma ISO 9001:2015 en el desarrollo ágil de software

María Fernanda Burdino+, Carlos Salgado+, Mario Peralta+, Alberto Sánchez+, Álvaro Ruiz de Mendarozqueta*

+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
fburdino@gmail.com {csalgado, mperalta, alfanego}@unsl.edu.ar
*Departamento Sistemas de Información - UTN FRC (Facultad Regional Córdoba)
Grupo Lidicalso (Laboratorio de investigación y desarrollo en ingeniería y calidad del software)
aruizdemendarozqueta@gmail.com

RESUMEN

En la actualidad cada vez más las organizaciones de desarrollo de software buscan un factor diferenciador, que les permita ofrecer a sus clientes productos de calidad, reduciendo sus costos e incrementando su productividad. La implementación y certificación de sus sistemas de gestión de Calidad bajo la norma ISO 9001:2015 (ISO ORG, 2015), es considerada una opción para lograr sus objetivos, pero se plantea la necesidad de lograrlos utilizando metodologías ágiles.
El presente trabajo tiene como objetivo desarrollar una guía para la implementación de la Norma ISO 9001:2015 (ISO ORG, 2015) utilizando la filosofía y métodos ágiles en organizaciones de desarrollo de software.
Se pretende investigar la implementación combinada de metodologías ágiles y modelos de calidad, identificar alternativas basadas en métodos ágiles a los requerimientos de implementación planteados en la guía ISO 90003 (ISO ORG), relacionar las buenas prácticas mencionadas en la guía de implementación ISO/IEC/IEEE 90003 (ISO ORG) con las prácticas ágiles implementadas en diferentes empresas del mercado del software.

Palabras Claves: Guía ISO/IEC/IEEE 90003, Norma ISO 9001, Metodologías Ágiles, Calidad de Software.

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 Nº 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. Se trabaja en lo referente a distintos modelos, métodos de evaluación y guías o estrategias de calidad con la participación del Grupo de Investigación "LIDICALSO" (Laboratorio de investigación y desarrollo en ingeniería y calidad del software) perteneciente a la UTN Facultad Regional Córdoba.

1. INTRODUCCION

Debido a la fuerte competencia en el mundo del desarrollo de software actual, las empresas y organizaciones dedicadas a esta actividad, buscan un factor que las diferencie de sus competidores, permitiéndoles ofrecer a sus clientes productos de calidad, reduciendo sus costos e incrementando su productividad. Para ello, muchas empresas intentan certificar calidad en sus procesos de producción. Así, la implementación y certificación de sus sistemas de gestión de Calidad bajo la norma ISO 9001:2015 (ISO ORG, 2015), es

considerada una opción para lograr dichos objetivos.
En cuanto al proceso de desarrollo, en el campo de la ingeniería de software. Existen distintas metodologías para su implementación. Dentro de ellas, se destacan las metodologías ágiles. Dichas metodologías, se caracterizan por adaptar la forma de trabajo a las condiciones del proyecto, brindando una mayor flexibilidad y una respuesta inmediata para acondicionar el proyecto y su desarrollo a las circunstancias específicas que se requieran.
Una de las metodologías ágiles más utilizadas es SCRUM.
Scrum es un framework de trabajo, que permite a las organizaciones implementar mejor gestión en sus proyectos de desarrollo de software, como así también incluir prácticas, técnicas y procesos que la organización considere importantes para entregar un producto a tiempo y de calidad a sus clientes. Está basado en valores y principios y es fundamental que los mismos sean respetados por cada organización o equipo que implemente la metodología SCRUM (ScrumAlliance).
El compromiso, el coraje, la concentración, la mente abierta y el respeto son los valores que los equipos SCRUM (Scrum Alliance, 2017) deben incluir en su trabajo diario hacia el éxito. Los principios expresados en el Manifiesto Ágil (agilemanifesto.org) son los siguientes:

- Satisfacción del cliente, mediante entregas tempranas y continuas
- Aceptación de cambios en los requisitos.
- Entregas de software funcional en el período más corto posible
- Equipo de desarrollo, trabajando junto a los responsables de negocio
- Equipo motivado, trabajando en un entorno adecuado
- Comunicación constante y fluida en el equipo de proyecto
- Software funcionando
- Desarrollo sostenible
- Excelencia técnica y buen diseño
- Simplicidad
- Auto organización del equipo
- Reflexiones del equipo para ser más efectivo

El framework de SCRUM contiene, en las actividades que plantea para la gestión de un proyecto, estos principios, los cuales están relacionados con los principios en los que se basa la Norma ISO 9001:2015 (Mendarozqueta, Principios y filosofía ágil).
SCRUM establece prácticas, las cuales son llevadas a cabo por sus respectivos roles (el Propietario del producto (Producto Owner), SCRUM Master y el equipo de desarrollo (Team Development)); los cuales, ejecutando ciertas actividades (Planificación del Sprint, reunión diaria, Ejecución del Sprint, Revisión del Sprint, Reunión Retrospectiva, Refinamiento del Backlog) en un lapso de entre 2 a 4 semanas obtienen un producto de calidad y funcionando, el cual si el Propietario del producto lo desea podría llegar a ser utilizado (Scrum Alliance, 2017).

SCRUM es un framework que es utilizado para entregar productos al cliente con mayor valor agregado, calidad, bajos costos, otorgando transparencia en la gestión y el desarrollo de los mismos tanto para el cliente como para el equipo y los altos mandos de la organización.

SCRUM según la encuesta desarrollada por la SCRUM ALLIANCE (ScrumAlliance) en organizaciones de diversos países del mundo, es el framework que tanto solo como en combinación con otros modelos es el más utilizado en las organizaciones a nivel mundial. El 16% de las organizaciones encuestadas utiliza solo SCRUM y el 78% utiliza SCRUM combinándolo con otros modelos como KANBAN; Cascada o Lean. (Scrum Alliance, 2018)

Respecto de las normas de calidad, la norma ISO 9001 (ISO ORG, 2015) se basa en un conjunto de principios establecidos para la correcta implementación de un Sistema de Gestión de Calidad en una organización que produce productos u ofrece servicios.

Los principios se desarrollan en diez capítulos normativos, los cuales establecen comportamientos o conductas esperables de una organización, las cuales permitirán satisfacer al cliente, entregando productos y/o servicios de calidad.

La implementación de la norma ISO 9001 (ISO ORG, 2015) permite entregar productos y servicios que respeten los requisitos establecidos por el cliente, logrando la satisfacción del mismo. Además permite gestionar riesgos e identificar oportunidades, las cuales permitan captar nuevos mercados.

En la encuesta realizada por la CESSI (CESSI, 2017) al evaluar las certificaciones de calidad que poseen las empresas del sector, la encuesta manifiesta que al menos un 67% de las empresas tenía algún tipo de certificación a diciembre de 2017 y el 66% del total tenía al menos ISO9001. Por otra parte cuando las empresas fueron consultadas por sus objetivos de I+D, el 41% de las mismas manifestó que pretende mejorar la calidad de sus productos.

Otro aporte importante en este ámbito es la guía ISO/IEC/IEEE 90003 (ISO ORG). Esta es una guía que puede ser utilizada por las organizaciones que desarrollan productos de software y pretenden certificar ISO 9001 (ISO ORG, 2015).

La guía provee información sobre prácticas para implementar los puntos normativos en la industria del software, utilizando lenguaje específico del desarrollo del software.

Sin embargo, la guía no especifica cómo estas prácticas deben implementarse, esto es trabajo de la organización, encontrar la manera adecuada de realizarlo. Además, la guía se basa en las metodologías de desarrollo de software tradicionales, no brindando información respecto a las prácticas ágiles de SCRUM (ScrumAlliance); Kanban (Wester) o Lean (LEAN Org).

Bajo estas consideraciones, se está trabajando en la implementación combinada de metodologías ágiles y modelos de calidad, identificar alternativas basadas en métodos ágiles a los requerimientos de implementación planteados en la guía ISO/IEC/IEEE 90003 (ISO ORG), relacionar las buenas prácticas mencionadas en la guía de implementación ISO/IEC/IEEE 90003 (ISO ORG) con las prácticas ágiles implementadas en diferentes empresas del mercado del software.

2. LINEAS DE INVESTIGACION Y DESARROLLO

Los ejes del tema de investigación son:

- Metodologías ágiles (Scrum, Kanban y Lean) **¡Error! No se encuentra el origen de la referencia.**
- Implementación combinada de metodologías ágiles y modelos de calidad.
- Alternativas basadas en métodos ágiles a los requerimientos de implementación planteados en la guía ISO/IEC/IEEE 90003.
- Relación de las buenas prácticas mencionadas en la guía de implementación ISO/IEC/IEEE 90003 con las prácticas ágiles.
- Factibilidad de implementación de las prácticas ágiles propuestas en diferentes empresas del mercado del software.

3. RESULTADOS OBTENIDOS/ESPERADOS

La Norma ISO/IEC/IEEE 90003:2018, es una guía para la aplicación de la norma ISO 9001:2015 (ISO ORG, 2015) en el desarrollo, implementación y mantenimiento de software, en organizaciones que desarrollen software. Desde este punto de vista, el objetivo general de la presente línea de investigación es complementar la guía ISO/IEC/IEEE 90003:2018, para las empresas de desarrollo que utilizan metodologías ágiles, específicamente SCRUM (ScrumAlliance).

Si bien existen trabajos de investigación relacionados, los mismos fueron basados en la anterior versión de la norma ISO 9001. (Mendarozqueta & Oliva, UCASAL, 2016). Estos trabajos son tomados como base para la presente investigación, teniendo en cuenta la trazabilidad planteada por ISO entre los

puntos normativos de la versión 2008 con la versión 2015.

A partir del año pasado las empresas que deseen certificar ISO 9001 lo deben realizar implementando sus Sistemas de Gestión de Calidad basándose en la versión 2015, es por ello que surge la necesidad de evaluar y validar nuevamente la compatibilidad de las metodologías ágiles y la norma ISO 9001 (ISO ORG, 2015), como así también incorporar el aporte de las prácticas ágiles en la ISO/IEC/IEEE 90003 (ISO ORG).

Se ha confeccionado una matriz, que vincula los puntos normativos de la ISO 9001:2015 (ISO ORG, 2015), la ISO/IEC/IEEE 90003:2018 (ISO ORG) y las prácticas ágiles del modelo SCRUM (ScrumAlliance). La base de la matriz es la relación de los principios de la gestión de calidad planteados en la norma ISO 9001:2015 con los principios planteados en el manifiesto ágil. La siguiente tabla muestra la relación:

Principios de la Gestión de Calidad (ISO 9001)	**Principios Manifiesto Ágil**	**Observaciones**
Enfoque al cliente	Satisfacción del cliente, mediante entregas tempranas y continuas	Tanto la Norma ISO 9001 como SCRUM están comprometidos con la satisfacción del cliente.
Liderazgo	Equipo de desarrollo, trabajando junto a los responsables de negocio Equipo motivado, trabajando en un entorno adecuado Comunicación constante y fluída en el equipo de proyecto Auto organización del equipo	Para que un equipo pueda trabajar de manera eficiente y efectiva, debe contar con buenos líderes organizacionales, compromiso de todos los involucrados hacia el objetivo y una comunicación constante. SCRUM es un modelo que promueve en su naturaleza estos principios.
Compromiso de las personas		
Gestión de las relaciones		
Mejora	Aceptación de cambios en los requisitos. Reflexiones del equipo para ser más efectivo	Los cambios son bien aceptados en cualquier equipo SCRUM, es por ello que permite a medida que se avanza con el desarrollo, mejorar el producto. También el marco de trabajo de SCRUM permite reflexionar al equipo sobre su desempeño, lo cual ayuda a mejorar para el siguiente Sprint.
Toma de decisiones basada en la evidencia	Entregas de software funcional en el período más corto posible Software funcionando Desarrollo sostenible Excelencia técnica y buen diseño	SCRUM marca un proceso de trabajo, el cual es repetible, basándose en información diaria para la toma de decisiones. Incorporando prácticas de ingeniería de software, permite desarrollar
Enfoque a procesos		

	Simplicidad	productos los cuales satisfagan a los clientes tanto por la calidad como por el cumplimiento en las fechas de entrega acordadas.

4. FORMACION DE RECURSOS HUMANOS

En relación a la presente línea de investigación, se está trabajando en el Proyecto de Tesis de la Maestría en Calidad de Software (Plan Ord. 017/09-CD) de la Ing. María Fernanda Burdino, de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis. Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Ciencias de la Computación y trabajos finales en el marco de la Ingeniería en Informática.

5. BIBLIOGRAFIA

agilemanifesto.org. (s.f.). *Principios Agiles.* Recuperado el 2016, de http://agilemanifesto.org/iso/es/principles.html.

CESSI. (2017). *Reporte anual del sector de software y servicios informáticos de la República Argentina.* Recuperado el 01 de 03 de 2019, de http://www.cessi.org.ar/descarga-institucionales-2219/documento2-02e5557d4fc5b9af48a726500a8c5bdd

ISO ORG. (Septiembre de 2015). *ISO 9001:2015.* Recuperado el octubre de 2016, de http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm

ISO ORG. (s.f.). *ISO 90003.* Recuperado el 2018, de https://www.iso.org/standard/74348.html

LEAN Org. (s.f.). *LEAN*. Recuperado el 12 de 26 de 2016, de http://www.lean.org/WhatsLean/

Mendarozqueta, A. R. (s.f.). *Principios y filosofía ágil.* Obtenido de http://www.slideshare.net/AlvaroRuizdeMendaroz/principios-giles

Mendarozqueta, A. R., & Andriano, N. (2014). *Un enfoque para la mejora continua basado en los principios ágiles.* Argentina: JAIIO.

Mendarozqueta, A. R., & Oliva, P. M. (30 de 12 de 2016). *UCASAL*. Recuperado el 2019, de http://bibliotecas.ucasal.edu.ar/opac_css/index.php?lvl=notice_display&id=61382

Scrum Alliance. (2017). *Scrum Guide.* Recuperado el 2019, de https://www.scrumalliance.org/why-scrum/scrum-guide

Scrum Alliance. (2018). *State of Scrum Report.* Recuperado el 2019, de https://www.scrumalliance.org/ScrumRedesignDEVSite/media/ScrumAllianceMedia/Files%20and%20PDFs/State%20of%20Scrum/2017-SoSR-Final-Version-(Pages).pdf

Scrum Primer. (s.f.). Recuperado el 2016, de http://www.scrumprimer.org/

ScrumAlliance. (s.f.). *ScrumAlliance.* Recuperado el 2019, de www.scrumalliance.org

Wester, J. (s.f.). *What is Kanban*. Recuperado el 2016, de http://www.everydaykanban.com/

Incorporar análisis de dominio (JODA) al proceso de desarrollo de UML components (DSBC)

Moyano Ezequiel UNTDF, Urciuolo Adriana UNTDF, Gel Matías UNTDF, Iturraspe Rodolfo UNTDF, Villarreal Martín UNTDF
Instituto de Desarrollo Económico e Innovación, UNTDF

Dir.: Fuegia Basket 251, (9410) Ushuaia. Tierra del Fuego. Tel: ++54-2901-432403
jmoyano@untdf.edu.ar,aurciuolo@untdf.edu.ar, mgel@untdf.edu.ar, riturraspe@untdf.edu.ar, mvillarreal@untdf.edu.ar

Resumen

El desarrollo de software basado en componentes se ha convertido en uno de los mecanismos más efectivos para la construcción de grandes sistemas, sobre todo en aquellos de gran complejidad. Construir una aplicación se convierte en la búsqueda y ensamblaje de componentes. UML Components presenta un gran potencial para construir sistemas basados en componentes en dominios complejos, y permite definir una manera de modelar.

El proceso de desarrollo UML Components debe adaptarse a los fines de proveer una mejor definición del dominio, para representar la complejidad del mismo. Se ve la necesidad de completar la etapa de modelado de requerimientos propuesta, incorporando el desarrollo de una etapa de Análisis de Dominio en forma previa.

El propósito del Análisis de Dominio (AD) es el de proporcionar la reutilización de la especificación de un dominio específico para aplicaciones similares. Una de las técnicas de análisis de dominio que mejor se ajusta a sistemas complejos es JODA (Object-Oriented Domain Analysis Method)*, ya que trabaja con análisis orientado a objetos, utilizando notación UML.*

El presente trabajo incorpora la técnica JODA (Análisis de Dominio) como primer etapa al proceso de desarrollo de software basado en componentes UML Components, a los fines de obtener consistencia en el proceso de modelado.

Palabras clave: Desarrollo de software basado en componentes, Análisis de dominio, Sistemas complejos, Reuso.

Contexto

La línea de investigación se desarrolla en el Instituto de Desarrollo Económico e Innovación (IDEI) de la UNTDF, por parte de un Grupo de docentes-investigadores que lleva adelante proyectos en la temática de Sistemas de Información Ambiental – Hidroinformática desde hace más de 10 años (acreditados por la UNPSJB previamente a la creación de la UNTDF).

Actualmente esta línea se desarrolla en el marco de un proyecto de investigación (Abr/2017-Mar/2019) – en Sistemas de Información Ambiental. El mismo ha obtenido aval UNTDF.

El Proyecto se lleva adelante con un enfoque multidisciplinario por parte de organismos de investigación y gestión en la provincia de Tierra del Fuego y consta de diversos componentes.

Introducción

Muchos dominios específicos, como los sistemas de información ambiental, se caracterizan por su complejidad [07], rozan una gran cantidad de subdominios diversos del conocimiento. [05]

Los trabajos de investigación actuales vinculados a esta problemática, no se han detenido en el estudio de técnicas de análisis de dominio apropiadas para comprender la complejidad y sus efectos, razón principal por

cual no se cuenta con modelos del dominio que puedan reutilizarse para distintos sistemas.

La ingeniería de dominio tiene como objetivo lograr la reutilización del conocimiento y de las especificaciones de un dominio específico.[09]

El Análisis del Dominio(AD) es introducido por Neighbors para estudiar el problema de un dominio o un conjunto de aplicaciones.[03] Las técnicas de AD están asociadas a la reutilización, su principal característica es capturar información relacionada con el dominio, y determinar qué y cómo este conocimiento se reutilizará en el desarrollo de futuras aplicaciones. El AD produce un *modelo del dominio*. [04]

El análisis de dominio se constituye en el punto de partida para obtener software reusable, sobre todo para sistemas de gran complejidad, facilitando y permitiendo el reuso de los productos obtenidos para las actividades futuras del desarrollo de software. (Fig. 1)

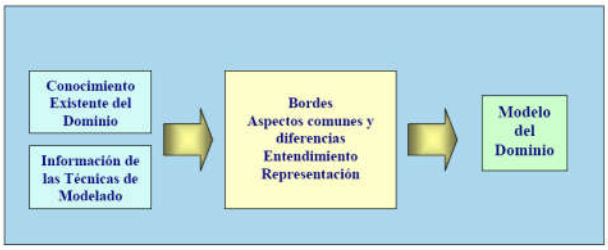


Fig 1. Análisis de Dominio

El Análisis de Dominio está enfocado en la reutilización de componentes de software dentro de un conjunto de sistemas que pertenecen a un dominio.[09]

Por otra parte la necesidad de contar con nuevos sistemas que no demanden grandes periodos de tiempo y esfuerzos (humanos y económicos), motivó la idea de contar con componentes ya desarrollados que permitieran su reutilización. Esto favoreció el avance de lo que se conoce como Desarrollo de Software Basado en Componentes (CSBD).[01]

El DSBC se ha convertido actualmente en uno de los mecanismos más efectivos para la construcción de grandes sistemas y aplicaciones de software.

Se constituye en un paradigma que permite el desarrollo de sistemas que presentan gran complejidad y magnitud, y brinda soporte para la integración de partes de sistemas mayores facilitando una estructura de ensamblado adecuada.[08]

Muchos son los beneficios que brinda el DSBC y que a su vez propiciaron la aparición de numerosas técnicas, herramientas y procesos de desarrollo, para la construcción de software basado en componentes.

El DSBC se define en diferentes etapas, entre ellas las propuestas por RUP (Rational Unified Process) en la cual se identifican cuatro etapas, o las propuestas por el proceso de desarrollo "UML Components", en éste caso el proceso se divide en seis etapas.

Si bien hay que tener en cuenta que la producción de software basado en componentes implica una gran cantidad de problemas de variada diversidad (que aún son temas de interés e investigación) existen actualmente tecnologías fundamentales que permiten abordar el problema de la complejidad, como el Análisis de Dominio, los cuales constituyen el tema central de investigación.

Las particularidades que poseen estos tipos de sistemas, promueven que para construir componentes verdaderamente reusables en el dominio, se necesiten definir técnicas apropiadas.

El presente trabajo centra su interés en el Análisis de Dominio y su aplicación en las etapas tempranas de los procesos de Desarrollo de Software Basados en Componentes, como punto de partida para el desarrollo de sistemas que presentan gran complejidad, que permita definir componentes verdaderamente reusables.

Para lo cual se utilizara la técnica de AD JODA incorporada a UML Components como proceso de DSBC.

La contribución principal de JODA es la separación de la ingeniería del dominio de la ingeniería de aplicación y su interacción. JODA pone el énfasis en el negocio y planeamiento de la metodología como punto de partida de la ingeniería del dominio.[03]

Es un método comprensivo "puro" de análisis de dominio y puede ser directamente aplicado, si se utiliza el método de análisis orientado a objetos (OOA). La meta de JODA es definir un modelo del dominio que se pueda utilizar para producir objetos reutilizables del software satisfaciendo sus requerimientos [10]. Se escogió JODA, entre otros, por su versatilidad y dado que trabaja con análisis orientado a objetos y notación UML, a los fines de obtener consistencia en el proceso de modelado DSBC.

UML Components es un proceso de desarrollo basado en componentes, utiliza diagramas de UML y estereotipos para extender UML[02].

El propósito de este método es el desarrollo de software basado en componentes, ya que se establece que son más adaptables y menos cambiantes que los objetos, utiliza diferentes herramientas produciendo componentes de software verdaderamente reutilizables. Se basa en un modelo conceptual de la información que existe en el dominio del problema, que es relevante al alcance del sistema modelado.[06]

Su objetivo principal es capturar los conceptos e identificar las relaciones.

El Proceso UML Component abarca desde la especificación de los requerimientos hasta el ensamblaje, donde cada proceso presenta los objetivos principales. El uso de UML para modelar todos los componentes para una solución integradora, da una visión constante de la solución entera y sus beneficios son muy significativos.

Líneas de Investigación y Desarrollo

La línea de investigación general se enfoca en el estudio y desarrollo para incorporar técnicas de análisis de dominio al proceso de desarrollo basado en componentes para dominios complejos. Para tal fin se utilizará la técnica de Análisis de dominio JODA al Proceso UML Components en su primer estapa.

JODA divide el proceso del análisis del dominio en tres fases [03]:

- Preparación del dominio.
- Definición del dominio.
- Modelado del dominio.

La preparación del dominio: produce dos productos, el material y la ayuda del dominio por parte de expertos del dominio.[13]

La definición del dominio: enumera todo el material fuente, referencias, sistemas, y expertos del dominio usados para el análisis.

Modelado del dominio: refina la definición del dominio y produce el modelo del dominio, consiste en diversos diagramas. Esta fase agrega términos al glosario del dominio.

Uno de los principales problemas observados en los procesos de DSBC, es que las técnicas de especificación y modelado de requerimientos que proveen no contemplan métodos detallados de análisis del dominio. La metodología UML Components pone el foco en la especificación, no en el Análisis de Dominio.

Debido a esto, se ve la necesidad de completar la etapa de modelado de requerimientos propuesta por UML Components, incorporando técnicas de análisis de dominio (JODA) en forma previa. Esto va a permitir una mejor definición del dominio y lograr comprender sus aspectos comunes, con el propósito de representar el conocimiento y permitir su reutilización.

Se redefine el proceso de desarrollo UML Components para extender su utilización a sistemas complejos, como son los sistemas de información ambiental que representa nuestro caso de estudio (Manejo sostenible de Cuencas)

En primer lugar se redefine el proceso de Requerimientos de UML Components,[13] se agrega la construcción de un modelo de dominio que constituya la entrada a la

definición del modelo de requerimientos en el proceso. Esto se refleja en las figuras 2 y 3.

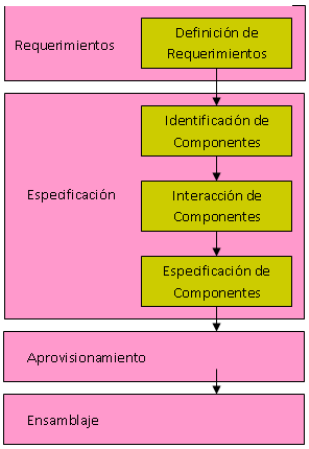


Fig. 2 Proceso UML Components.

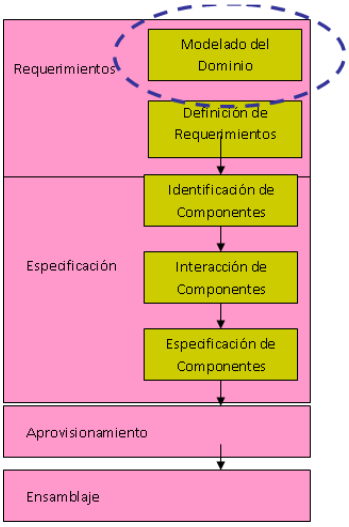


Fig. 3 Redefinición del Proceso UML Components.

Al incorporar el modelado del dominio como entrada a los requerimientos en el proceso de UML components, se debe a su vez redefinir el proceso de especificación de requerimientos.

Para eso se deben considerar las etapas definidas en la técnica JODA para el modelado del dominio.

La figura 4 muestra a continuación como queda definida la nueva fase para especificación de requerimientos en UML Components, incorporando el modelado de dominio definido en JODA[13]

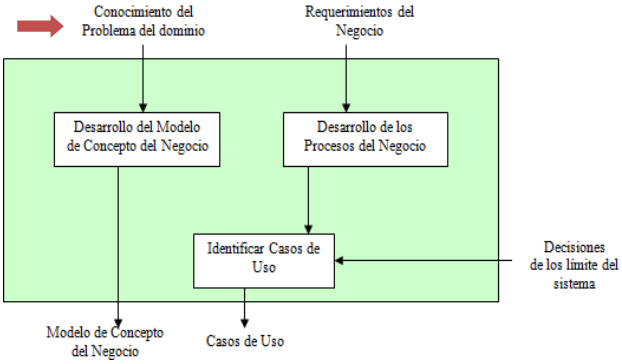


Fig. 4 Definición de Requerimientos UML Components.

Como se observa, lo que en el modelo se llama "Conocimiento del problema del dominio", se formaliza mediante la producción de los distintos diagramas contemplados en JODA para cada una de las fases que propone en el análisis del dominio [13]. Fig 5.

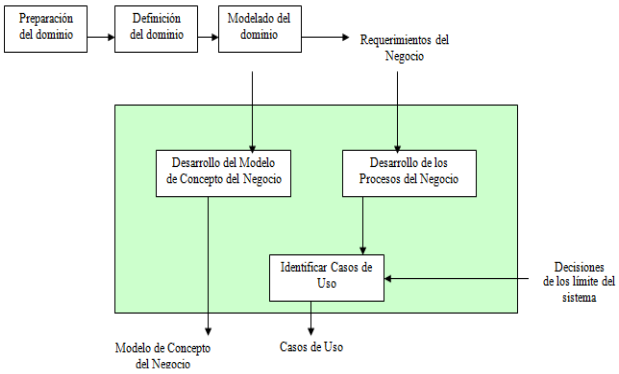


Fig. 5 Definición de Requerimientos UML Components.

A continuación se propone un modelo basado en UML para la etapa de modelado del dominio que fue incluida al Proceso de definición de Requerimientos, según las fases que define JODA para tal fin. Se utiliza la notación UML para brindar consistencia a todo el proceso de desarrollo UML Components. El siguiente cuadro muestra la propuesta para representar el Modelo del Dominio en UML.[13]

FASES DE JODA	Notación JODA	Notación UML
Preparación del Dominio	Modelo Informal	Se mantiene el modelo Informal
	Servicios del dominio	Diagrama de Works Units

Definición del Dominio	Diagrama de Contexto Diagrama Conjunto-Parte	Se utiliza diagrama de Packages con actores
	Dependencias del dominio	Se refleja en el diagrama de Dependencias entre Packages
Modelado del Dominio	Diagrama de Clases de Coad y Yourdon	Diagrama de Clases de nivel conceptual de UML
El diagrama de Procesos del Dominio se muestra en el diagrama de Actividades que porpone de UML Components		

Resultados

Si bien la línea de investigación es incipiente, el grupo de investigación ha obtenido resultados de proyectos que constituyen un insumo para las actuales investigaciones.

A partir de aplicar Análisis de Dominio a UML Components permitió obtener un proceso de desarrollo (sobre todo en la etapa de especificación de requerimientos) más detallado, preciso y formalizado, extendiendo UML Components a sistemas de mayor complejidad.

Utilizar la notación UML en la técnica JODA, permitió definir un número de modelos y diagramas para la especificación del software con varios grados de abstracción, con lo cual se obtuvo un modelado del dominio más semejante a la realidad.

Como trabajos futuros se propone validar el resultado del presente trabajo, en un dominio específico y obtener componentes reusables, en al menos dos aplicaciones del dominio.

Formación de Recursos Humanos

El Equipo de Trabajo está conformado por docentes investigadores de la UNTDF y expertos del dominio. En ésta línea de investigación se encuentran en desarrollo dos tesis de posgrado (Magister en Ing. de Software, UNLP, Ezequiel Moyano y Martín Villarreal).

Colaborará y prestará apoyo un alumno en formación de grado, con el objetivo formarlo y que le sirva para el desarrollo de su futura tesis de grado.

Referencias

[01] Manuel F. Bertoa, José M. Troya Y Antonio Vallecillo *Aspectos De Calidad En El Desarrollo De Software Basado En Componentes, 2002.*

[02] Cheesman, John; Daniels John, *UML Components,* Octubre 2000.

[03] Robert Holibaugh *Joint Integratedavionics Working Group (JIAWG) Object-Oriented Domain Analysis Method (JODA), 1994.*

[04] Maarit Harsu. *A Survey On Domain Engineering, Institute Of Software SystemsTampere University Of Technology.*

[05] Arne Koschel, Ralf Kramer, Ralf Nikolai, *A Federation Architecture For An Environmental Information System Incorporating GIS, Karlsruhe, Germany, 2000.*

[06] Elsa Estévez, *Desarrollo De Software Basado En Componentes,* Neuquen 2003.

[07] Günther, O. Springer-Verlag*, Environmental Information Systems, Berlín Heidelberg, Germany, 1998.*

[08] Luis Iribarne Martíne*, Un Modelo De Mediación para el Desarrollo de Software Basado en Componentes COTS, Universidad De Málaga, Julio 2003.*

[09] Vanessa Hamar *Aspectos Metodológicos Desarrollo y Reutilización de Componentes de Software, Mérida – 2003.*

[10] Emeline, Regis; Gustavo, Tondello; *Análise de Domínio,(Universidad e Federal de Santa Catarina, 2000)*

[11] Urciuolo Adriana, Iturraspe Rodolfo, Moyano Ezequiel. *Perfil UML 2.0 para Aplicaciones de Monitoreo Ambiental.* VI JIISIC'07, Lima, Perú. Facultad de Cs e Ing., Pontificia Universidad Católica del Perú 2007, ISBN 978-9972-2885-1-7. pp. 393-401.

[12] Urciuolo A., R. Iturraspe, *"Conceptual Patterns for Water Resources Information Systems*", En: *Journal of Computer Science and Technology* Vol. 3 - No. 1 - April 2003 - ISSN: 1666-6038, pp 20-26.

[13] Moyano Ezequiel *Tesis de Grado: Técnicas de Análisis de Dominio para el Desarrollo de Componentes en Sistemas Complejos, Caso de Estudio: Sistemas de Información Ambiental, dirigida por* Urciuolo A. 2007.

Ingeniería de Software Dirigida por Modelos aplicada a sistemas robóticos usando los estándares de la OMG

Claudia Pons, Gabriela Pérez, Roxana Giandini, Carlos Neil, Marcelo de Vincenzi

Comisión de Investigaciones Científicas de la Provincia de Buenos Aries (CIC)
Universidad Abierta Interamericana (UAI)
Facultad de Informática de la UNLP

Buenos Aires, Argentina

Resumen

El Desarrollo de software Dirigido por Modelos (MDD, Model Driven software Development) aparece como una alternativa viable para aplicar técnicas de ingeniería de software en el desarrollo de sistemas robóticos. Su uso logra un nivel de abstracción superior, permitiendo utilizar los estándares propuestos para robótica, y así obtener ventajas como generalidad, reutilización, claridad, expresividad. Estas son todas cualidades inherentes a un proceso de creación de software eficiente y eficaz. El objetivo general de esta investigación es contribuir al mejoramiento de los procesos de desarrollo de software de los sistemas robóticos, a través del análisis del paradigma de desarrollo MDD aplicando los estándares definidos por la OMG.

Palabras clave: sistemas robóticos, ingeniería de software, desarrollo dirigido por modelos, estándares de la OMG.

1. CONTEXTO

Los sistemas robóticos (RSS Robotic Software Systems) desempeñan un papel cada vez más importante en nuestra vida cotidiana. La necesidad de sistemas robóticos en todos los entornos (industriales, educativos) aumenta y sus requisitos se vuelven más exigentes. Están hechos de diferentes componentes y sensores, lo que resulta en una arquitectura muy compleja y altamente variable. Actualmente, la mayoría de los sistemas de software de robótica se basan todavía en software propietario y están estrechamente ligados al hardware específico, las plataformas de procesamiento o la infraestructura de comunicación. En consecuencia, estos robots solo pueden ser armados, configurados y programados por expertos. Los enfoques tradicionales utilizados en el proceso de desarrollo de este tipo de sistemas están basados principalmente en codificar las aplicaciones sin ningún tipo de técnica de modelado. Y, aunque estas aplicaciones se utilizan en diferentes sistemas robóticos, se pueden identificar algunos problemas. Entre ellos, vale la pena mencionar que no hay documentación clara sobre las decisiones de diseño que se toman durante la fase de codificación, por lo que se dificultan tanto la evolución como el mantenimiento de estos sistemas. Además, al utilizar lenguajes de programación específicos perdemos la posibilidad de generalizar conceptos que pueden ser extraídos, reutilizados y aplicados

en otros sistemas, lo que nos permitiría evitarnos recodificar todo desde cero cuando se lo necesite.

Por otro lado, a medida que los sistemas robóticos crecen para ser cada vez más complejos, la necesidad de aplicar los principios de ingeniería de software para su proceso de desarrollo se convierte en un reto obligatorio en estos días.

Desde esta perspectiva, se acepta el hecho de establecer nuevos enfoques para satisfacer las necesidades del proceso de desarrollo de sistemas robóticos tan complejos como los de hoy. El desarrollo basado en componentes (component-based development o CDB), la arquitectura orientada a servicios (Service-oriented architecture o SOA) [5], así como la Ingeniería de software dirigida por modelos (MDE), y el modelado específico de dominio (DSM) son algunas de las tecnologías más prometedoras en el dominio de los sistemas robóticos.

Actualmente se promueve la integración de los componentes de los sistemas robóticos a través de la adopción de estándares de la OMG [1, 11]. Se han lanzado cuatro especificaciones: para la interacción en los sistemas robóticos (Robotic Interaction Service – ROIs) [3], para la localización de los sistemas robóticos (Robotic Localization Service - RLS), para el modelado de los componentes (Robotic Technology Component - RTC) [2] y para el despliegue dinámico y su configuración (Dynamic Deployment and Configuration for Robotic Technology Component - DDC4RTC).

Es en este contexto en el que el Desarrollo Dirigido por Modelos (MDD, Model Driven Development)[7, 8, 9] aparece como una alternativa viable para aplicar técnicas de ingeniería de software en el desarrollo de este tipo de sistemas. Su uso logra un nivel de abstracción superior, permitiendo utilizar los estándares propuestos para robótica, y así obtener ventajas como generalidad, reutilización, claridad, expresividad. Estas son todas cualidades inherentes a un proceso de creación de software eficiente y eficaz.

2. OBJECTIVOS, HIPÓTESIS Y RESULTADOS ESPERADOS

El objetivo general de esta investigación es contribuir al mejoramiento de los procesos de desarrollo de software de los sistemas robóticos, a través del análisis del paradigma de desarrollo MDD aplicando los estándares definidos por la OMG para el modelado de estos sistemas y complementado con componentes (CBD) y la arquitectura orientada a servicios (SOA).

3. LINEAS DE INVESTIGACIÓN

En este proyecto convergen dos líneas de investigación: Robótica y MDE.

La problemática descripta arriba ha llevado a la concepción de un conjunto de actividades con el objetivo de acercar a los investigadores de estas aéreas: Ingeniería de Software dirigida por Modelos por un lado, y Robótica por el otro. Ejemplos de estas actividades son el Workshop on Domain-Specific Languages and Models for Robotic Systems (DSLRob) lanzado en 2009, y el Workshop on Model-Driven Robot Software Engineering(MORSE) iniciado en 2013, ambos con el objetivo de incentivar la interacción de estas aéreas. Como resultado, en los últimos años varias plataformas han sido desarrolladas para proveer formas simples e intuitivas de construir las aplicaciones de software robótico. Esto incluye tanto prototipos académicos como así también productos comerciales. Paralelamente el organismo OMG que se ocupa de coordinar los estándares de software internacionalmente ha creado el

grupo Robotic Domain task Force, para ocuparse de estos temas.

En el ámbito industrial/comercial las plataformas para desarrollo robótico más conocidas son: Lego Mindstorms Evolution 3, Choregraphe, Robotino View 2 y Microsoft Robotics Developer Studio 4 (MRDS4).

En el ámbito académico hemos encontrado muchas propuestas teóricas que aplican el paradigma CBD para el desarrollo de sistemas robóticos. Otras propuestas han aplicado SOA para construir estos sistemas. También el paradigma MDD/MDA/DSL ha sido aplicado de manera teórica a la construcción de robots (ver nuestro survey en [21]).

Todos estos trabajos proveen un punto de partida robusto para el desarrollo de nuestro proyecto.

4. PROPUESTA

Este trabajo tiene como objetivo general estudiar los sistemas robóticos y cómo aplicar técnicas de ingeniería de software para desarrollarlos. Este objetivo está siendo abordado a través de los siguientes sub-objetivos:

- Estudiar las propuestas de la OMG para el modelado de sistemas robóticos, y analizar como pueden ser aplicadas en el desarrollo de los sistemas robóticos.
- Analizar las interrelaciones entre los distintos estándares.
- Analizar cómo pueden ser integrados los paradigmas SOA y CBD con lo propuesto por la OMG
- Estudiar como los modelos realizados en dichos estándares pueden ser transformados para crear otros modelos, y eventualmente, crear código.

5. METODOLOGÍA DE INVESTIGACIÓN Y DESARROLLO PROPUESTA

A partir del estudio de los sistemas robóticos actuales, se observa que son artesanales, y que su desarrollo esta basado principalmente en el código. A pesar de que existen propuestas para el modelado de los sistemas robóticos, aun no se aplican y tampoco cubren la posibilidad de la integración de dichas propuestas.
Nuestra hipótesis de trabajo consiste en la posibilidad de integrar los estándares con otras técnicas de ingeniería de software para poder aplicar MDD en el desarrollo de los sistemas robóticos.
Respecto a los lenguajes de modelado y transformación nos basaremos en los estándares de la OMG, en particular MOF [6], UML, OCL y QVT[10].
Finalmente, todos nuestros resultados teóricos serán plasmados en herramientas de desarrollo de software de código abierto, preferentemente sobre la plataforma Eclipse [15].

6. ESQUEMA DE PLAN DE TRABAJO C/ACTIVIDADES

El proyecto viene desarrollándose a través de las siguientes actividades:

- Estudio de los lenguajes estándares para modelado de sistemas robóticos
- Estudio del estándar como base para el modelado de software, en particular MOF (Meta Object Facilities), RTC (Robotic Technology Component), ROIs (Robotic Interaction Service), DDC4RTC (Dynamic Deployment and Configuration for Robotic Technology Component), RLS (Robotic Localization Service).
- Estudio del estándar para modelado de servicios SOA.

_ Analizar las relaciones entre dichos estándares y como se complementan para ser utilizados en el proceso de desarrollo.
_ Analizar mecanismos de transformación entre modelos
_ Plasmar lo analizado en herramientas de desarrollo de software de código abierto
_ Evaluación de la propuesta mediante la aplicación de la misma en proyectos reales.

7. FORMACIÓN DE RECURSOS HUMANOS Y RESULTADOS

El equipo está integrado por cinco investigadores senior, provenientes de 2 universidades (UNLP y UAI), quienes combinan sus conocimientos en robótica e ingeniería de software. El resto del equipo está integrado por estudiantes de pregrado y postgrado. En el contexto de este proyecto se están desarrollando 5 tesis de licenciatura en Informática, 5 tesis de Maestría en TI y 5 tesis doctorales (1 financiada por CONICET y 1 por el programa DoctorAR). Se trabaja además en colaboración con la Universidad de Viena a través de un convenio bi-lateral financiado por el Mincyt [16]. Los resultados preliminares del proyecto se han publicado en [17], [18], [19] , [20] , [21] y [22].

8. REFERENCIAS

1. OMG Robotics-DTF - Object Management Group - http://robotics.omg.org/ (consultado en 2015)
2. Documento de especificación RTC 1.0 - Object Management Group - http://www.omg.org/spec/RTC/1.0/ (consultado en 2015)
3. Documento de especificación Robotic Interaction Service (RoIS) – OMG - http://www.omg.org/spec/RoIS/ (consultado en 2015)
4. Robotic Localization Service (RLS) – OMG - http://www.omg.org/spec/RLS/
5. Documento de especificación SOA – OMG - http://www.omg.org/technology/readingroom/SOA.htm (consultado en 2015).
6. Documento de especificación OMG's MetaObject Facility (MOF) Home Page - http://www.omg.org/mof/
7. Stahl, M Voelter. Model Driven Software Development. John Wiley, ISBN 0470025700.
8. Kleppe, Anneke G. and Warmer Jos, and Bast, Wim. MDA Explained: The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2003.
9. Object Management Group, MDA Guide, v1.0.1, omg/03-06-01 (2003).
10. MOF 2.0 Query/View/Transformations - OMG Adopted Specification. March 2005. http://www.omg.org.
11. Object Management Group (OMG) http://www.omg.org
12. Ledeczi, A., Bakay, A., Maroti, M., Volgyesi, P., Nordstrom, G., Sprinkle, J., Karsai, G. Composing Domain-Specific Design Environments. IEEE Computer 34 (2001)
13. Kleppe, Anneke. MCC: A Model Transformation Environment. A. Rensink and J. Warmer (Eds.): ECMDA-FA 2006, LNCS 4066, pp. 173 – 187, Spain, June 2006.
14. Atlas Model Weaver Project Web Page. http://www.eclipse.org/gmt/amw/, 2005.
15. Proyecto Eclipse - https://eclipse.org/ (consultado en 2015).
16. Project Title: “Adoxx Meta-Model Compiler. Modelling Methods for Robotic Systems”. Programa de

Cooperación Científico-Tecnológica entre el Ministerio de Ciencia, Tecnología e Innovación Productiva de la República Argentina (MINCyT) y el Ministerio de Ciencia e Investigación de la República de Austria (BMWF).

17. C. Pons, G. Pérez, R. Giandini, G. Baum . A Model-Driven Approach to Constructing Robotic Systems. Journal of Computer Science & Technology. Vol. 14 - No. 1 – April 2014 - ISSN 1666-6038.
18. Jerónimo Irazábal and Claudia Pons. Metamodel independence in Domain Specific Modeling Languages. "Communications in Computer and Information Science CCIS Series" vol.411, "Software and Data Technology" pp.140-154, ISBN 978-3-642-45403-5. Springer-Verlag Berlin Heidelberg (2013).
19. Omar Martinez Grassi, Claudia Pons, Gabriel Baum. Variable-Based Analysis for Traceability in QVT-R Model Transformations. CIbSE - SET 2015 (CIbSE -Software Engineering Track). XVIII Conferencia Iberoamericana en "Software Engineering", CIbSE 2015. Perú. April 2015.
20. Gabriela Pérez, Jerónimo Irazábal, Claudia Pons y Roxana Giandini. Applying MDE tools to defining domain specific languages for model management. SADIO Electronic Journal of Informatics and Operations Research. ISSN 1514-6774 vol. 12, no. 1 (Sept 2013).
21. Claudia Pons, Roxana Giandini, Gabriela Arévalo . A systematic review of applying modern software engineering techniques to developing robotic systems.. Vol. 32 No. 1 de 2012. Revista Ingeniería e Investigación. Tri-annual ISSN: 0120-5609. Indexada en el ISI.
22. Claudia Pons, Gabriela Pérez, Roxana Giandini and Gabriel Baum: "Applying MDA and OMG Robotic Specification for Developing Robotic Systems".Published in Lecture Notes in Computer Science. Springer. Volume 9959 2016. System Analysis and Modeling. Technology-Specific Aspects of Models. Saint-Malo, France, October 3-4, 2016. Editors:Jens Grabowski, Steffen Herbold ISBN: 978-3-319-46612-5 (Print) 978-3-319-46613-2 (Online)

|Ingeniería de software para sistemas embebidos, sistemas críticos y entrega continua de software

Irrazábal, Emanuel; Mascheroni, Maximiliano Agustín; Sambrana, Iván; Alonso, José Manuel ; Acevedo, Joaquín; Lezcano, Andrea

1: Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura. Universidad Nacional del Nordeste

eirrazabal@exa.unne.edu.ar

Resumen

Esta línea de investigación aborda temas de ingeniería del software en sistemas tradicionales y su traslado tanto a sistemas embebidos como a sistemas críticos. 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 soporte para el desarrollo de sistemas críticos con aplicación al ámbito ferroviario nacional.

Finalmente se está trabajando en el desarrollo de procedimientos para la formalización de las pruebas continuas, en la disciplina de entrega continua, lo cual facilitará los ensayos en los desarrollos ágiles.

Palabras clave: Ingeniería de software, entrega continua, sistemas embebidos, sistemas críticos

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.

Introducción

Ingeniería de software para sistemas embebidos y sistemas críticos

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.

Asimismo, en el marco de los sistemas embebidos se encuentran los sistemas críticos cuyo mal funcionamiento o fallo pueden resultar en pérdida de vidas humanas, equipos, instalaciones o daños medio ambientales severos [9]. Por ello, uno de los ámbitos de mayor utilización de sistemas críticos es el ferroviario.

En particular, 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 [3]. Sin embargo, todos los sistemas electrónicos para la seguridad vial de trenes y subtes son importados y muy caros. Esta situación ha favorecido que ocurran terribles accidentes [5] 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 [6][7] . 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.

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 [8]. 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 [9]: 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 [10]: 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 [11]: 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.

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, teniendo en cuenta el cumplimiento de la normativa EN 50128.

Ingeniería de Software para Entrega Continua

Tal y como lo describe la segunda ley de Lehman [12], 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 [13], hace a la validación de los requerimientos una etapa crítica en el desarrollo de las aplicaciones [14].

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 [15]. 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 [16][17].

Entonces, haciendo foco en la gestión de los requerimientos, es posible adaptar técnicas ágiles las cuales tendrán un fuerte impacto positivo [18][19], 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 [20].

Desde otro punto de vista, en el ámbito de fábricas de software se 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 [21]. El objetivo es poder lanzar a producción un producto software libre de defectos "*con solo apretar un botón*" [22].

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 y Sistemas Críticos:

- Desarrollar soluciones de sistemas embebidos con microcontroladores que soluciones problemas regionales y apoyen a los grupos de investigación locales.
- Realizar pruebas de concepto para la construcción de arquitecturas software y diseños específicos con patrones orientados a sistemas críticos de altas prestaciones.

En la línea de Ingeniería de Software y Entrega Continua:

- Utilizar el modelo de procedimientos para la gestión de requerimientos en PYMEs yerbateras a partir de técnicas ágiles.
- Desarrollar un modelo de diseño y ejecución de pruebas continuas de software, especialmente orientado a flaky test y crossbrowsing testing.

Resultados obtenidos

El grupo de investigación es de reciente formación, por lo cual se enumeran los resultados de este último año. A continuación se indican:

En la línea de Ingeniería de Software para Sistemas Embebidos y Sistemas Críticos:

- En [23] se realizó una revisión sistemática de la literatura con el fin de identificar las principales técnicas y

herramientas con las que sumar aspectos de seguridad a la gestión de requerimientos software tradicional.

- En [24] se describe un procedimiento para el diseño de arquitectura software orientada a sistemas críticos ferroviarios de acuerdo con la normativa EN 50128.
- En [25] se describe una metodología de desarrollo de software para aplicaciones ferroviarios de acuerdo con la normativa EN 50128.
- En [26] y [27] se describen en detalle dos implementaciones de proyectos software para sistemas críticos ferroviarios nacionales que se realizan con la metodología descripta en[25].

Asimismo, se realizó la primera Escuela de Mecatrónica Aplicada de Corrientes en el marco de un Acuerdo de Trabajo entre la Facultad de Ciencias Exactas y Naturales y Agrimensura (FACENA- UNNE) y el Ministerio de Industria, Trabajo y Comercio de la provincia de Corrientes [28].

En la línea de Ingeniería de Software para la Entrega Continua:

- En [29] se identificaron el conjunto de problemas que afectan a la calidad del software en las disciplinas de entrega continua y pruebas continuas teniendo en cuenta lo obtenido en revisiones sistemáticas de la literatura publicadas en años anteriores .
- En [30] se identificaron los factores claves para el éxito al realizar pruebas tipo *flaky* en servicios REST.
- En [31] se realizó una revisión sistemática de la literatura obteniendo las soluciones ante los problemas en pruebas continuas.
- En [32] también se realizó una revisión sistemática de la literatura orientada al cross-browser test.

Formación de recursos humanos

En el Grupo de Investigación en Innovación en Software y Sistemas Computacionales (GIISSC) están involucrados 5 docentes investigadores, 1 becario de investigación de pregrado, 1 tesista de doctorado y 3 tesistas de maestría. Cuatro 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] Sitio web con estadísticas CNRT: https://www.cnrt.gob.ar/content/estadisticas, Visitado: 19/02/2018.

[4] Ingeniería y gestión del mantenimiento en el sector ferroviario (Spanish Edition) Paperback – 2010, Arques Paton José Luis.

[5] Sitio web con ejemplos de accidentes ferroviarios argentinos:: https://es.wikipedia.org/wiki/Categor%C3%ADa:Accidentes_ferroviarios_en_Argentina. Visitado: 19/02/2018,

[6] Sitio web com ejemplo de licitación: https://www.clarin.com/ciudades/trenes-china-compra-licitacion-reestatizacion_0_rJduN7cwXe.html. Visitado:19/02/2018.

[7] https://www.clarin.com/ieco/china-trenes_de_carga-randazzo-inversiones_0_rJygK8mKP7l.html

[8] J. L. Boulanger, “CENELEC 50128 and IEC 62279 Standards”, Control, Systems and Industrial Engineering Series, John Wiley & Sons, Inc., 2015, p. 13.

[9] EN 50126. Railway applications - The specification and demonstration of Reliability, Availability, Maintainability and Safety (RAMS). 2005.

[10] EN 50128. Railway applications - Communication, signalling and processing systems - Software for railway control and protection systems. 2011.

[11] EN 50129. Railway applications - Communication, signalling and processing systems - Safety related electronic systems for signaling. 2005.

[12] 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.

[13] 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.
[14] G. Kotonya and I. Sommerville, Requirements engineering. Chichester: John Wiley & Sons, 1998.
[15] Plan Estratégico para el Sector Yerbatero – Yerba Mate Argentina", Yerbamateargentina.org.ar, 2016.
[16] 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.
[17] 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.
[18] 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", CICN 2014.
[19] One, V. "State of agile development survey results", 2017.
[20] 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.
[21] L. Chen, "Continuous Delivery: Huge Benefits, but Challenges Too" in IEEE Software 03/2015. V. 32(2).
[22] J. Humble and D. Farley. "Continuous delivery: reliable software releases through build, test, and deployment automation", 1st ed. Boston, US: Pearson Education, 2010.
[23] Pinto Luft, Cristian, Emanuel Irrazabal, and Iván Sambrana. "Revisión Sistemática de la Literatura: aplicación de seguridad a requerimientos software de sistemas críticos ferroviarios." XIX Simposio Argentino de Ingeniería de Software (ASSE)-JAIIO 47 (CABA, 2018). 2018.
[24] Irrazábal, Emanuel, Iván Sambrana, and Cristian Pinto Luft. "Proceso para el diseño de la arquitectura software en un sistema crítico ferroviario según la norma EN-50128." *XIX Simposio Argentino de Ingeniería de Software (ASSE)-JAIIO 47 (CABA, 2018)*. 2018.
[25] Irrazábal, Emanuel, Iván Sambrana, Cristian Pinto Luft, María de los Ángeles Gómez López, Sergio Gallina, Adrián Laiuppa, José Brizuela, Pablo Gomez, Ariel Lutenberg. "Metodología de desarrollo de aplicaciones ferroviarias según las normas ISO 9001 – EN 50128 " CASE 2018, Córdoba.
[26] Adrián Laiuppa, Martín Amado, José Manuel Cruz, Facundo Larosa, Irrazábal, Emanuel, Iván Sambrana, Sergio Gallina, María de los Ángeles Gómez López, Ariel Lutenberg. "Sistema automático para ensayos de ciclo de vida de relés ferroviarios de seguridad." CASE 2018, Córdoba.
[27] Adrián Laiuppa, Martín Amado, Sergio Gallina, Irrazábal, Emanuel, Iván Sambrana, Hugo Ferrari, Dario Baliña, Leandro Francucci, María de los Ángeles Gómez López, Juan Manuel Cruz, Pablo Gomez, Ariel Lutenberg. "Sistema de monitoreo remoto de barreras ferroviarias automáticas". CASE 2018, Córdoba.
[28] Acuerdo de Cooperación entre la FACENA y el Ministerio de Industria, Trabajo y Comercio de Corrientes para la puesta en marcha de la Escuela de Mecatrónica Aplicada de Corrientes según Res. 0539/18 CD Facena.
[29] Mascheroni, Maximiliano; Irrazábal, Emanuel. Problemas que afectan a la Calidad de Software en Entrega Continua y Pruebas Continuas. CACIC 2018
[30] Mascheroni, M. A., & Irrazábal, E. (2018). Identifying key success factors in stopping flaky tests in automated REST service testing. Journal of Computer Science and Technology, 18(02), e16-e16.
[31] Mascheroni, M. A., & Irrazábal, E. (2018). Continuous Testing and solutions for testing problems in Continuous Delivery: A Systematic Literature Review. Computación y Sistemas, 22(3).
[32] J Sabaren, L. N., Mascheroni, M. A., Greiner, C. L., & Irrazábal, E. (2018). A Systematic Literature Review in Cross-browser Testing. Journal of Computer Science and Technology, 18(01), e03.

Integración Universidad y Gobierno para una Efectiva Transformación Digital en el Ámbito de la Ingeniería de Software

Vilallonga, Gabriel[1,2]; Riesco, Daniel[1], Berón, Mario[1]; Sánchez, Alejandro[1]; Acosta Parra, Carlos[2] ; Sacaba, Flavio[2].

[1]Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis. Ejercito de los Andes 950. San Luis
Tel: +54 (0266) 4520300. Int. 2127
[2]Departamento Sistemas de la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca
Maximio Victoria Nº 55 - C.P: 4700 - San Fernando del Valle de Catamarca
Teléfono: 03834- 435112 – int 168

gvilallo@tecno.unca.edu.ar/unsl.edu.ar, driesco@unsl.edu.ar, mberon@unsl.edu.ar, asanchez@unsl.edu.ar, carlosacostap@tecno.unca.edu.ar, flaviomariano1993@gmail.com

Resumen

Las organizaciones actuales evidencian una creciente complejidad de sus procesos lo que conlleva a la necesidad de mejorar sus capacidades digitales que habiliten a un mejor ambiente de trabajo y servicios.

La línea de investigación que se presenta propone el estudio de la conformación de equipos de trabajo para el logro de una efectiva transformación digital desde una perspectiva de reingeniería de software.

La adecuada aplicación de este tipo de reingeniería debe contextualizarse en un marco más general y sinérgico, como lo es la reingeniería de procesos.

En ámbitos gubernamentales, en la mayoría de los casos, no cuentan con los procesos de los organismos. Esto hace aún más ardua la aplicación de reingeniería en dichos organismos.

En dicho contexto es que se plantea como eje central la formación de equipos de trabajo mixtos, personas con conocimientos de la organización interna de las reparticiones gubernamentales en conjunto con profesionales informáticos con formación en ingeniería de software.

En esta propuesta se plantea la hipótesis de que la conformación de estos tipos de equipos permiten una efectiva reingeniería de software sobre una solidad reingeniería de procesos lo que habilita a una implantación rápida y efectiva de sistemas de información colaborando efectivamente en la transformación digital deseada.

Para llevar a cabo este trabajo se ha establecido ejes de capacitación en organismos gubernamentales al mismo momento que alumnos de carrera de Ingeniería Informática cursan la materia Reingeniería de Proceso y Sistemas de Información. La primera tarea es capacitar los recursos humanos de los entes en notaciones estándar de procesos para un efectivo relevado de los procesos, junto con el contexto en reingeniería en que está inserta este tipo de herramienta.

Esto permite obtener un detallado conocimiento de los procesos y su ámbito de implantación provisto por el personal de dichas reparticiones. En esta tarea el aporte de la visión informática asiste al correcto análisis de los sistemas informáticos que son utilizados.

A partir de la reingeniería de los procesos relevados permite una propuesta de mejora de los procesos con especial énfasis en el análisis de los sistemas de información que forman parte de dichos proceso. Esto habilita a la detección de debilidades de los sistemas y propuestas de mejoras de los mismos. Es en este punto se reducen los tiempos de implementación de las mejoras por la interacción de los dos grupos de personas que componen el equipo de trabajo.

Este es un aporte importante a la vinculación de la universidad con el medio y de las prácticas de los alumnos en ámbitos reales de aplicación de los conocimientos recibidos.

Palabras Claves: Ingeniería de Software. Reingeniería de Procesos. Sistemas de Información. Transformación Digital. Investigación y Consultoría.

Contexto

El proyecto de investigación "Fortalecimiento de la Calidad y la Productividad en Ingeniería de Software" es evaluado por la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Catamarca y se desarrolla y ejecuta en ámbitos del Departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas de la UNCa y en el Departamento de Informática de la Facultad de Ciencias Físico, Matemáticas y Naturales de la Universidad Nacional de San Luis donde los integrantes se desempeñan como docentes de la carrera de Ingeniería en Informática (UNCa), e Ingeniería en Informática y Licenciatura en Ciencias de la Computación (UNSL) respectivamente.
Esta línea de investigación, inserta en el proyecto, establece como prioridad la formación científico-técnico de los integrantes con la premisa de proceder a la transferencia de resultados a los medios donde la informática juega un rol fundamental en la transformación digital.

A inicios del año 2018, se celebró un convenio marco de cooperación entre la Facultad de Tecnología y Ciencias Aplicadas de la Universidad Nacional de Catamarca y la Municipalidad de la Ciudad de San Fernando del Valle Catamarca. Este convenio prevé el trabajo conjunto entre la Dirección de Modernización de la Administración Municipal y la Cátedra de Reingeniería de Procesos y Sistema de Información de la Carrera Ingeniería en Informática de la casa de altos estudios.
En 2019 se ha establecido un nuevo convenio entre los mismos participantes, FTyCA de la UNCa y el municipio. En esta ocasión se prevé desarrollar trabajos similares en Rentas Municipal de la municipalidad de SFV de Catamarca.
La cooperación ha tenido y tiene como objetivo mejorar la calidad de los procesos por medio de su TD, que incremente las capacidades de estos. Una de las formas de lograrlo ha sido la transferencia de recientes investigaciones llevadas a cabo en el área de ingeniería de software de nuestra facultad. Unas de las tareas han sido la formación de los recursos humanos del ente municipal y de los alumnos de la carrera de Ingeniería en Informática en lo que respecta a la apropiación de capacidades prácticas en medios reales de aplicación de los conocimientos.

Introducción

Desde el año 2010 la Facultad de Tecnología y Ciencias Aplicadas (FTyCA) de la Universidad Nacional de Catamarca (UNCa), de la República Argentina, creó el Laboratorio de Tecnologías de la Información y las Comunicaciones (LaTICs) con el fin de proveer a los docentes de su facultad un ámbito de desenvolvimiento científico, académico y de transferencia. Todos los proyectos de investigación de la unidad académica con perfiles tecnológicos se encuentran bajo su ámbito de acción. Uno de los fines estratégicos de la creación es "Promover la transferencia tecnológica y la asistencia técnica al medio". Basados en esta premisa se han llevado a cabo actividades de investigación, consultoría y/o transferencia a los distintos actores tecnológicos regionales. Dependiendo de las particularidades de las actividades estas son realizadas como servicios a terceros rentados, o no, atendiendo al tipo de vinculación que se haya establecido entre la Universidad y las entidades participantes.
Una de las actividades principales de este laboratorio ha sido la de llevar a cabo investigaciones acerca de la adecuación de "la teoría" con respecto a su aplicación práctica

referidos a Ingeniería de Software en general. Esta actividad de corroboración, llevado a cabo en entes gubernamentales, ha permitiendo transferir conocimientos adquiridos, por actividades de análisis y estudio de reingeniería de procesos, en estos tipos de organizaciones. Estos servicios han habilitados la retroalimentación entre los conceptos y prácticas referidas a la TD por medio de las tecnologías de la información y los ámbitos reales de aplicación.

En este trabajo se plantea una hipótesis acerca de la conformación de equipos de trabajos referidos a tareas tendientes a la reingeniería de software basada en la reingeniería de procesos.
El trabajo se plantea desde la perspectiva consultoría-transferencia, de parte de la universidad hacia el municipio, y prácticas profesionales supervisadas en ámbitos administrativos municipales llevados a cabo por los alumnos de la materia Reingeniería de Procesos y Sistemas de Información (RPySI) de la carrera Ingeniería Informática de la FTyCA, bajo la supervisión de los docentes de la materia.
Experiencias previas realizadas desde el proyecto sugieren equipos conformados en forma mixta entre los actores. La experiencia permite establecer la siguiente hipótesis de trabajo: "La efectiva TD debe ser abordada desde un grupo mixto de trabajo que habiliten a una transformación digital en el menor tiempo posible. Una parte del grupo es especialista en el negocio con amplio conocimiento del manejo interno del área, especialmente provisto por el ente. El resto de los participantes es un grupo técnico con capacidad de mejora de procesos por medio de implantación de tecnologías de la información, como por ejemplo el desarrollo de distintos sistemas informáticos."
Esta estructura dentro de la organización permite acelerar notablemente la TD de la organización, al menos parcialmente, dependiendo del grado de calidad de digitalización de sus procesos.
Bajo esta estructura de equipo de trabajo se intenta mostrar en forma efectiva y mensurable resultados satisfactorios en diferentes procesos que en un comienzo presenten baja calidad de implementación, con subutilización de recursos humanos y pobre gestión documental.
La metodología de trabajo propuesta permite un rápido feedback ya que al conformar este equipo de trabajo entre alumnos y agentes de la repartición se produce la capacitación en paralelo en aspectos de reingeniería lo cual son inmediatamente aplicados en un medio real como lo es Rentas Municipal, mejorando la experiencia de 2018 en la Dirección de Modernización.

Se debe capacitar en forma adecuada, desde lo conceptual, del concepto de proceso y de contar con un manejo apropiado de herramientas de diagramas de procesos para, una vez identificado los procesos, proceder al análisis y rediseño. Para esto se provee de técnicas de relevado de procesos y su posterior notación por medio de una notación de procesos de negocios (BPMN, de su sigla en inglés), para una adecuada documentación.
El relevado de los procesos permite conocer en detalle aquellos procedimientos que presentaban baja calidad de implementación.
El trabajo de determinación de calidad de procesos está basado en Factores Críticos de Éxito (FCE). Esto permite establecer métricas referidas a la calidad de implementación con sustento en la matriz de ponderación de cada uno de los procesos con respecto a los FCE.
El trabajo por medio de esta técnica permite establecer con claridad el orden de prioridad de los procesos con mayor urgencia de rediseño para lograr un rápido impacto en la mejora parcial de los objetivos plantados. Esta tarea, sin la asistencia de personas amplio conocimiento de la estructura organizativa, es ardua. Al contar con un equipo con el conocimiento necesario allana el camino hacia una mejora de los procesos ineficientes.

El trabajo consta de una sección dedicada brevemente a los procesos y su notación. En forma seguida se da una breve descripción del trabajo con FCE en la medición de calidad de los procesos. Una sección aborda, en forma

escueta, la detección de los procesos ineficientes por medio de matrices especialmente formuladas para contar con sustento que justifica el trabajo sobre un proceso antes que otros en lo que refiere a la aplicación de reingeniería.

Procesos y Notación de Procesos de Negocios

Un proceso en nuestro trabajo, es una colección de tareas relacionadas y estructuradas que describen como se realiza, o se produce, un determinado servicio o producto [1]. Usualmente se visualizan como un flujo de trabajo que establecen las reglas del negocio. Además, el flujo de trabajo puede ser esquematizado en el ámbito de diferentes estructuras de departamentos, permitiendo, en etapas de análisis, detectar y evitar los silos funcionales.

La Gestión de Procesos de Negocios (BPM, de su sigla en inglés) utiliza distintos métodos para incrementar la calidad de los procesos de negocios. En estas prácticas es necesario contar con cada proceso especificado para su revisión continua.

La BPMN es una herramienta que permite la diagramación de procesos de negocios la cual se usa comúnmente en ambientes académicos y comerciales. Esta forma de notar ha sido desarrollada para que en forma sencilla y, en períodos cortos, integrantes de proyectos de reingeniería con perfiles diferentes posean un lenguaje común de comunicación. Dicha notación gráfica estándar permite diagramar los procesos en forma de flujo de trabajo. Esta notación cuenta con el soporte de la Object Management Group (OMG). En [2] se provee una guía rápida de cada uno de los elementos de la notación.

Reingeniería de Procesos: Factores Críticos de Éxitos y la Gestión de Calidad de los Procesos

En ingeniería de software los indicadores de calidad son necesarios para contar con métricas que permitan manejar indicadores de calidad de implementación de los procesos relevados. Un problema común en reingeniería es por donde comenzar el análisis de los mismos para tener producir cambios visibles en tiempos cortos. Una forma de establecer la calidad de implementación es los trabajos basados en FCE en el análisis de los procesos.

Los FCE son objetivos claves que deben alcanzarse para cumplir una misión. En los organismos gubernamentales argentinos no están definidos, por lo que se debe trabajar en su definición para permiten medir la calidad de los procesos, especialmente en servicios.

Esta forma de trabajo fue inicialmente propuesta en [3]. Luego en [4] se presenta metodología para la implementación de estrategias administrativas, el cual es una adaptación de lo presentado en [3] y que provee una guía para consultores.

Estas metodologías han sido adaptadas a nuestra necesidad de métricas de calidad de implementación, y por la urgencia de contar con resultados rápidos de reingeniería y por consiguiente la mejora en la organización, o parte de esta.

Identificación de Procesos Ineficientes

Se utilizan dos técnicas, “Matriz de Procesos / FCEs” y la “Matriz de FCEs de un Proceso / Valoración de la Calidad de Implementación del Proceso”. En la primera se ordenan los procesos según la prioridad otorgada por los directivos del área. En la segunda se detecta sobre cuál de los procesos se debe comenzar a trabajar primero. Los procesos con mala calidad de implementación y con una gran cantidad de FCEs asociados serán tratados primeros para obtener un rápido impacto en la reestructura de las funcionalidades del área.

Para una descripción detallada se sugiere consultar [5].

Línea de Investigación, Desarrollo, e Innovación

Inserto en el marco del proyecto se encuentra la línea de investigación referida la revisión continua de conceptos y prácticas de Ingeniería de Software efectivamente realizada en ámbitos reales de aplicación. La FTyCA de la UNCa promueve estos tipos de

transferencias con doble beneficio, la transferencia al medio y la formación de alumnos en áreas de desarrollo de software y de definición y análisis de procesos.
Esta línea de investigación permite abrir una perspectiva multidisciplinar que tiene como efecto sinérgico el trabajo con ámbitos reales con recursos humanos técnicos no informáticos que enriquecen la disciplina.
Estas actividades han exigido realización de cursos acerca de la temática, como así también el estudio de material bibliográfico y de publicaciones científicas.
Es de hacer notar que estos grupos de trabajos están coordinados por docentes investigadores de las distintas universidades.

Resultados y Objetivos

El objetivo principal de esta línea de investigación es el de contribuir de manera efectiva y mensurable en la concreción de aportes concretos al área de IS aplicada, lo que implica revisión, o nuevas propuestas, de técnicas, metodologías, que asistan al desarrollo de software inserto en procesos optimizados.
El efecto deseado, también, es incidir significativamente en las actividades académicas de grado y posgrado, la formación de recursos humanos. Estos objetivos están siendo alcanzado gracias al trabajo conjunto entre los equipo de la UNSL-UNCa.

Formación de Recursos Humanos

Esta línea de investigación está en una etapa intermedia, donde integrantes del proyecto desarrollan sus tesis de doctorado y maestría en el marco de la carrera de doctorado en Ingeniería de Software, como así también de la maestría en Ingeniería del Software en temas específicamente relacionados a la línea de Reingeniería de Procesos e Ingeniería de Software. Además los participantes pertenecen a distintas cátedras de las carreras de ingeniería y licenciatura de las universidades que participan.
Se ha procedido a la incorporación de alumnos de los últimos años con la finalidad de incluirlos en actividades de investigación y desarrollo en las áreas específicas del proyecto
El proyecto prevé un programa de capacitación y formación de recursos humanos, que contempla las siguientes actividades de dirección de tesinas de grado de la carrera de Ingeniería en Informática y en actividades de actualización y posgrado en el área de estudio.
Se propone la participación de los integrantes en eventos nacionales e internacionales de la especialidad, como congresos, simposios, seminarios y cursos.

Bibliografía

[1] Weske, M. (2012). «"Chapter 1: Introduction"». Business Process Management: Concepts, Languages, Architectures. Springer Science & Business Media. p. 1–24. ISBN 9783642286162.
[2] http://www.bpmn.org/-www.bpmnquickguide.com/view-bpmn-quick-guide/
[3] Hardaker, M., and B. K. Ward, “How To Make a Team Work,” Harvard Business Review, Vol. 65, No. 6, November–December 1987, pp. 112–119.
[4] Pinto, Juan De J. - ASIGNACIÓN Y DETERMINACIÓN DE PRIORIDADES DE PROCESOS ESENCIALES, CON BASE EN LOS FACTORES CRÍTICOS DE ÉXITO - Estudios Gerenciales, núm. 74, enero-marzo, 2000, pp. 79-89 Universidad ICESI Cali, Colombia - ISSN: 0123-5923.
[5] http://latics.tecno.unca.edu.ar/convenios/docs Convenios, Informes realizados en el presente trabajo.

Línea de Investigación en Evaluación de Productos *Software:* dificultades encontradas

Paula Angeleri, Andrés Blanco, Jorge Ceballos, Rolando Titiosky,
Facultad de Ingeniería y Tecnología Informática, Universidad de Belgrano, Argentina
{paula.angeleri; andres.blanco; jorge.ceballos, rolando.titiosky} @comunidad.ub.edu.ar

RESUMEN

El objetivo de este artículo es presentar las dificultades encontradas durante el año 2018, dentro de la Línea de investigación MyFEPS *Metodologías y Framework para la Evaluación de Productos de Software* [1]. 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 como un proyecto de investigación en Universidad de Belgrano, en el año 2010 para dar apoyo al IRAM[1] en su servicio de certificación de productos software [3]. En la fase inicial del proyecto se especificó un modelo de calidad más actualizado que los modelos publicados hasta ese momento, el modelo de calidad QSAT [4], y un proceso de evaluación que permite establecer una ponderación de ítems de calidad (características, sub-características, sub-sub-características, y atributos) en base a la importancia relativa que definen los *stakeholders*[2], considerando su percepción de los objetivos de la evaluación de producto software, el proceso de evaluación MyFEPS [5]. En [1][6][7][8][9][10]. se explican otros objetivos que tuvo el proyecto

1. INTRODUCCIÓN

MyFEPS se ha convertido en una línea de investigación, con un propósito principal, que es realizar validaciones y mejoras continuas del *Framework MyFEPS*, y sublíneas de investigación, entre ellas la principal se apoya en el proyecto MyFEPS2 [2], que busca ganar agilidad en la evaluación de productos software (sin perder calidad). En este contexto, en el último año cuyos resultados estamos presentando, se definió un proceso ágil de evaluación de productos software [2] y se realizaron algunos proyectos pilotos de evaluación ágil de productos software, tales como productos CRM. Como resultado de estos proyectos se han validado nuevas características del modelo de calidad QSAT [4], ampliado el alcance del método de evaluación MyFEPS [5], y actualizado la base de herramientas de apoyo que conforman el Framework MyFEPS [11].

Avances alcanzados

En [1] se explicaron 2 sub líneas de esta investigación y los avances de sendos proyectos que trabajaron esas sub líneas, una de ellas es la **Evaluación de motores para video juegos FPS, donde se llevaron a cabo los proyectos de evaluación de** tres *Game Engines*: Unity 3D, Unreal Engine y Cry Engine, la otra sublínea se orientó a la **Evaluación de Tecnologías de Backup,**

[1] Instituto Argentino de Normalización y Certificación.

[2] Partes interesadas en la evaluación, como pueden ser el Instituto de Certificación, el Laboratorio, el Desarrollador, y/o actores del proyecto que toman decisiones.

llevándose a cabo la evaluación del sistema de *Backup* Empresarial, "Avamar" de la empresa EMC, y el "producto TSM" de la empresa IBM.

En este *paper* se resumirá la experiencia de evaluar los productos CRM indicados a continuación:

Evaluación de la calidad de productos CRM:

El objetivo principal del proyecto de Evaluación de Herramientas CRM fue el de evaluar al menos 3 productos CRM (en su versión gratuita) con el fin de recomendar el que más se ajustaba a los objetivos del proyecto. Como resultado de las encuestas a los *stakeholders* del proyecto, se decidió evaluar las Características Básicas del Modelo de Calidad QSAT: Constancia, Efectividad, y Satisfacción subjetiva de los usuarios. Se analizaron las características no funcionales, cómo responden los tiempos de procesamiento de los productos CRM y sus parámetros de calidad al alterar la cantidad de usuarios en un entorno de testing, comparar el número de funciones implementadas con el número de funciones definidas en el documento de requerimientos y poder determinar el grado de satisfacción obtenido durante las pruebas.

Se estudiaron puntualmente tres productos CRM: **Zoho CRM, Free CRM y Hubspot CRM.**

El proceso de evaluación de la calidad utilizado fue MyFEPs Ágil, con el que se definieron las etapas, las estimaciones de costo, los tiempo necesarios y los recursos humanos y sus roles del proyecto.

El modelo de Calidad elegido fue el modelo QSAT al cual no fue necesario incorporar nuevas características de calidad, ni sub-características, ni atributos con sus métricas.

Durante el proyecto, se evaluaron las Características Básicas (CB) de QSAT: CB1 "Constancia", CB2 "Efectividad", CB3 "Satisfacción subjetiva de los usuarios".

Las subccaracterísticas evaluadas fueron CB1 SC1 "Independiente de la cantidad de información acumulada",

CB2 SC1 "Cobertura de funcionalidades útiles",

CB2 SC2 "Cumplimiento con las capacidades esperadas",

CB3 SC1 "Satisfacción total". Los atributos evaluados fueron

CB1 SC1 A1 "Tiempos de procesamiento con distintos números de usuarios en entorno de testing",

CB1 SC1 A2 "Empeoramiento de parámetros de calidad con distintos números de usuarios en entorno de testing"

CB2 SC1 A1 "Diferencia entre el número de funciones implementadas y el número de funciones definidas en el documento de requerimientos",

CB2 SC2 A2 "Diferencia entre el número de funciones implementadas y el número de funciones definidas en el documento de requerimientos",

CB3 SC1 A1 "Grado de Satisfacción total percibido en entorno de prueba".

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Como se explicó en [1], las líneas principales de investigación sobre las que se ha trabajado son: el Dominio de conceptos, el Modelo de calidad de producto software [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 influencian 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 tesinas y trabajos académicos.

En el TFC de D. Ardizzone se mejoró el formato del Informe de resultados de la evaluación en un desarrollo BI[18].

La sub-línea de investigación, abordada por los tesistas S. Christello [19], E. Romanin [20], y C. 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.

La sub línea de investigación abordada por el tesista 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 tesistas 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 tesista 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]. A partir de haber detectado la necesidad de resolver cuestiones de calidad de desarrollo vinculadas con la seguridad de la información en la web y en el marco de la integración académica entre materias, durante el año 2019 se pretende iniciar el estudio conjunto entre la materia "Ingeniería de Software IV, Testing y Calidad", con la materia "Seguridad Informática", para aplicar el modelo MyFEPS a los casos de desarrollo de software, utilizando el marco de OWASP como atributos para integrar con QSAT. De esta manera se podrían complementar los tópicos de la calidad con requisitos de la seguridad informática de una aplicación. Esto nos enfrenta a la posibilidad de extender el alcance de MyFEPS a otros planos.

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 [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 utilizó MyFEPS en diferentes escenarios, con diferentes propósitos, con distinta metodología de evaluación, con distinto contexto del producto evaluado (en desarrollo, en producción). 5. Se utilizó MyFEPS considerando particularidades especiales de software de diferentes ámbitos, como ser: productos e-commerce, videojuegos, CRMs, software académico, de la industria de la Salud, entre otros. 6. Se utilizó MyFEPS con diferentes stakeholders y 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, Sindicatura Gral de CABA, 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 pregrado, grado, en cursos de postgrado y de actualización profesional, en las universidades UB y UNLaM, así como también en la UTN, en la maestría de Ingeniería en Calidad, en el Seminario Inspección y Ensayos, dando herramientas a los alumnos para evaluar software de ensayos de laboratorio. En particular, durante el 2018 en la materia Ingeniería de Software IV Testing y Calidad de UB se destaca, entre otros, el proyecto de los alumnos I. Fernández Mendoza, T. Vukasovic Kippes, E. Moreno Villa y M. Nennhuber han podido realizar la Evaluación de conformidad del Software denominado “Blockchain” versionado 1.0 usando el Framework MyFEPS, con el modelo QSAT de calidad de software, evaluando la funcionalidad y la integridad de datos. Las características y sub-características evaluadas fueron: CB1 Calidad de los Artefactos: SC1.1 Modularidad; CB2: Correctitud: de Procesos; CB2 Efectividad: SC2.1 En relación a las funcionalidades; CB3 Instalabilidad: SC3.1 Primera instalación.

Los proyectos mencionados han servido para validar la *Capacidad de ser entendido* del Framework MyFEPS, su adecuación, y se han realizado ajustes al Modelo QSAT. En particular, en las últimas evaluaciones piloto se ha observado que el Framework era entendible para alumnos de 4to y 5to grado de UB, mientras que resultaba complejo para una becaria de 2do. año de Ingeniería en Informática, y confuso para alumnos de 4to. año de otra Universidad, con otro terminal como carrera tecnológica. A nivel posgrado siempre fue fácil de entender. A nivel tecnicatura, fue fácil de entender, pero esto puede deberse a que casi todos los alumnos tienen experiencia laboral. A nivel de cursos lo han entendido bien, posiblemente porque

los cursos eran muy orientados al tema. En 2019 se continuará trabajando para que el Framework sea fácilmente entendido por alumnos universitarios de carreras tecnológicas de todos los niveles. Otro aspecto que surge de los proyectos pilotos de validación del Framework es la necesidad de contar con ítems de calidad que tuvieran en cuenta aspectos económicos, cuando éste es un parámetro considerado en la decisión final.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo contó inicialmente con dos co-Directores, 2010 a 2012, y luego continuó con un solo Director. Participaron 3 profesores investigadores principales, y empleados de una empresa. Colaboraron 2 profesores investigadores de la UB, y un profesor de la PUCP, Perú. Participaron al menos 4 becarios, y 1 alumno de la PPS. Del 2012 a la fecha participaron 10 tesistas de grado y numerosos alumnos de UB (pregrado y grado), UNLaM (grado) y UTN (posgrado).

5. BIBLIOGRAFIA

[1] Angeleri, P., Titiosky, R; Ceballos J., Aguilera, S.; Línea de investigación en evaluación de productos software WICC2018 - ISSN: 978-987-3619-27-4, 731-736, 2018.

[2] Titiosky, R.; Angeleri, P., MyFEPS.Proceso de evaluación AGIL de productos de software v3, Documento de trabajo, Universidad de Belgrano, 08/06/2016.

[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, CONAIISI 2014, ISSN: 2346-9927 - Página 1043.

[5] Angeleri, P.; Titiosky, R.; Santi, M..; Davila, A.: Proceso de Evaluación de Productos *Software*, CONAIISI 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 - JUI 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, CONAIISI 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' CONAIISI, 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, Nº 72, 2015.

Medición de confianza interpersonal en equipos globales de desarrollo de software

Sergio Zapata[1], José Luis Barros-Justo[2], Estela Torres[1], Gustavo Sevilla[1], Facundo Gallardo[1]

[1]Instituto de Informática - FCEFN - Universidad Nacional de San Juan {szapata, etorres, gsevilla, fgallardo}@iinfo.unsj.edu.ar

[2]Escuela de Informática - Universidad de Vigo (España) jbarros@uvigo.es

RESUMEN

El progreso alcanzado por las tecnologías de información y comunicación ha promovido fuertemente las relaciones o vinculaciones entre personas de distintos puntos geográficos del mundo. Bajo este novedoso contexto nuevas configuraciones de equipos de desarrollo de software surgen, conocidos como equipos virtuales.

Un desafío en estos nuevos escenarios distribuidos es el entendimiento de los aspectos humanos y sociales involucrados en el proceso de desarrollo de software y su impacto en la efectividad del mismo.

El objetivo del presente trabajo es identificar y clasificar conocimiento reportado acerca de la medición de confianza interpersonal en contextos globales de desarrollo de software. Para alcanzar ese objetivo se aplica un estudio de mapeo sistemático de la literatura científica.

***Palabras clave*:** Ingeniería de Software Global, Confianza Interpersonal, Equipos Virtuales.

CONTEXTO

Este trabajo de investigación está enmarcado en el proyecto de investigación denominado "Colaboración, Confianza y Desarrollo Global de Software" (código 21/E1061), financiado por la Universidad Nacional de San Juan. Siendo la unidad ejecutora el Instituto de Informática (IdeI) de la misma Universidad.

El proyecto es de ejecución bianual y su comienzo fue en Enero de 2018. El mismo está inserto dentro del Programa de Incentivos a Docentes Investigadores de la Secretaría de Políticas Universitarias del Gobierno Nacional.

1. INTRODUCCIÓN

Los avances en Tecnologías de Información y Comunicación (TIC) han promovido vinculaciones, sociales y laborales, entre personas de diferentes zonas geográficas, creando así nuevos desafíos tecnológicos, culturales y organizacionales.

Las organizaciones han migrado a trabajo basado en equipo. Los avances de las TIC promueven y motivan esa migración. La conformación de quipos ha pasado de una configuración con personal del mismo lugar (co-localizados) hacia una configuración que incluye personal de distintas locaciones geográficas y de diferentes organizaciones [1].

El resultado es la creación de equipos virtuales con nuevas formas de trabajo, nuevos estilos de toma de decisión y diferentes formas de relaciones humanas. Estos cambios plantean desafíos teóricos, prácticos y de investigación sobre la efectividad de equipos de trabajo en estos nuevos contextos [1], [2], [3]. Especialmente debido a que los factores que impactan el desempeño de los equipos tradicionales no aplican o son menos efectivos en los equipos virtuales [1].

Esta tendencia también emerge en el negocio del desarrollo de software, es decir grupos de desarrolladores que trabajan geográficamente distribuidos.

Adicionalmente, mucha investigación se ha focalizado en el entendimiento de las emociones y de los estados de ánimo en el desarrollo de software y cómo estos aspectos humanos pueden afectar los resultados finales de una disciplina técnica como la Ingeniería de Software [4], [5], [6], [7], [8], [9], [10].

Así, un factor clave es la confianza interpersonal [11]. Esto es así en equipos co-localizados de desarrollo de software y es más relevante cuando estos equipos operan en un contexto virtual o global. [12].

Mayer et al [13] definen confianza como "la disponibilidad de una parte a ser vulnerable a las acciones de otra parte". La confianza permite a las personas participar en actividades riesgosas que ellas no pueden controlar o monitorear; y por lo tanto podrían ser perjudiciales para ellas mismas [14], [15].

Se cree que la confianza es un factor fundamental en la determinación del éxito o fracaso de un equipo virtual [16], [17], [18]. Dado que un alto nivel de confianza produce comunicación social significante, comunicación y feedback predecible, pensamos que la confianza es un prerrequisito para alcanzar la coordinación efectiva de los trabajos grupales [19].

Las investigaciones muestran que los equipos con alto grado de confianza son más proactivos, más centrados en los resultados de las tareas, más optimistas, con interacciones más frecuentes y con feedbacks más productivos [20].

El desarrollo de software es claramente una actividad humana y como tal es propensa a mejoras continuas de performance. La medición de software es un enfoque para controlar y administrar el proceso de software mediante el seguimiento y la mejora de su performance [21].

Las técnicas de medición de software promueven la mejora del control del proceso de desarrollo de software, reduce tiempos y costos de desarrollo; y producen software de más alta calidad [22]. Medidas de software han sido promovidas como recursos esenciales para mejorar la calidad y controlar el costo durante el desarrollo de software [23], [24]. Creemos que el nivel de confianza es un indicador valioso que puede servir para tomar decisiones en orden a mejorar la performance de un equipo virtual de desarrollo de software.

El propósito de este trabajo es identificar y clasificar conocimiento reportado acerca de la medición de confianza interpersonal en contextos globales de desarrollo de software.
Para alcanzar ese objetivo aplicaremos un estudio de mapeo sistemático (SMS por sus siglas en ingles, Systematic Mapping Studies) con el fin de buscar, recuperar y revisar la literatura científica; seleccionar estudios relevantes y extraer datos de interés. El proceso de ejecución del SMS es mostrado en la Fig.1.

Un SMS provee una estructura de reportes y resultados de investigación publicados mediante la categorización de ellos y a menudo dando un resumen visual, un mapa, de los resultados [25]. La intensión de un SMS, como un método de investigación, es identificar toda la evidencia científica relacionada a un tema específico [26].

Nuestro interés es identificar y clasificar evidencias acerca de técnicas de medición de confianza interpersonal, métricas de confianza y metodologías de desarrollo donde estas mediciones son aplicadas. Todo ello enmarcado en contextos GSD.

Para el presente SMS hemos establecido las siguientes preguntas e investigación:

RQ1: ¿Cuales técnicas de medición son usadas para medir confianza en GSD?
RQ2: ¿Cuales métricas son usadas para medir confianza en GSD?
RQ3: ¿Qué decisiones son tomadas en función de las mediciones de confianza en GSD?
RQ4: ¿Cuales metodologías de desarrollo de software son usadas en los procesos donde se mide confianza?

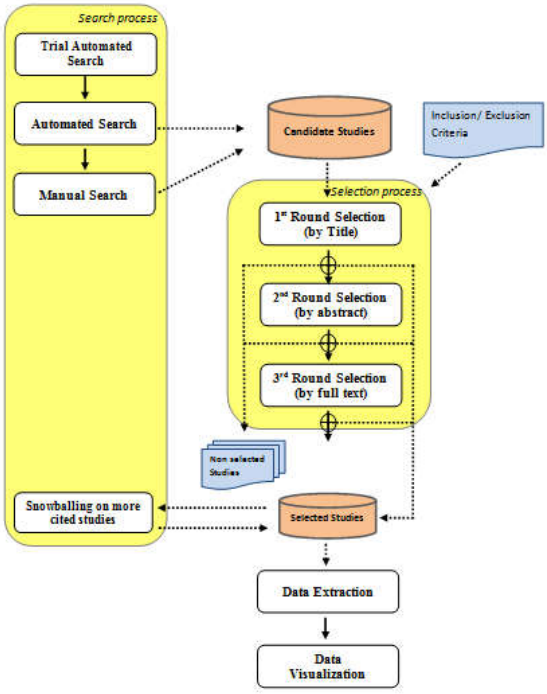


Fig. 1. Proceso de ejecución del SMS

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación sobre la cual se ejecuta el presente trabajo intenta generar conocimiento científico sobre los aspectos sociales involucrados en equipos distribuidos o globales de desarrollo de software. Seaman [27] argumenta que la Ingeniería de Software es una compleja disciplina pues combina tanto aspectos técnicos como aspectos humanos.

Cada vez es más reconocida la importancia que tienen los vínculos sociales de los miembros de los equipos de desarrollo y su impacto en la efectividad del proceso de software. La complejidad de entender estas relaciones humanas se incrementa cuando los miembros de esos equipos están distribuidos y las distancias físicas, temporales, culturales, de lenguaje afectan las comunicaciones e interacciones de los mismos.

Bajo este contexto de investigación es que se desenvuelven los actividades del presente trabajo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Hasta el momento se han obtenido los siguientes resultados del trabajo propuesto:

- Definición de un protocolo de SMS, en el cual se han detallado cada una de las actividades que se ejecutarán en el trabajo de revisión literaria. Participan investigadores de Argentina, Chile, Uruguay y España.
- Se han revisado aproximadamente 700 papers identificados mediante búsquedas automáticas, búsqueda manual y técnica de snowballing [28].
- Actualmente el proceso SMS se encuentra en la etapa de extracción de datos.

Se espera, en el corto plazo, contar con resultados valiosos que logren responder las preguntas de investigación planteadas inicialmente en el trabajo.

4. FORMACIÓN DE RECURSOS HUMANOS

Debido a la magnitud del trabajo, especialmente las actividades de revisión de la literatura, el equipo de trabajo lo conforman

una importante cantidad de personas. Además de los autores colaboran profesores de la Universidad ORT (Uruguay), prof. Gerardo Maturro, de la Universidad de Quindío (Colombia), prof. Faber Giraldo, y de la Universidad de la Frontera (Chile), prof. Samuel Sepúlveda.

El presente trabajo forma parte de una tesis de doctorado que se desarrolla en el Programa de Doctorado en Ciencias de la Informática que se dicta en la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan (UNSJ).

Algunos resultados de este trabajo serán insumo para una tesis de graduación de carrera de Licenciatura en Ciencias de la Computación que también se dicta en la UNSJ.

5. BIBLIOGRAFÍA

1. Berry, G. R. (2011), Enhancing effectiveness on virtual teams: Understanding why traditional team skills are insufficient. Journal of Business Communication, 0021943610397270.
2. Tannenbaum, S.I., Mathieu, J.E., Salas, E. and Cohen, D. (2012), "Teams are changing: are research and practice evolving fast enough?", Industrial and Organizational Psychology: Perspectives on Science and Practice, Vol. 5 No. 1, pp. 2-24.
3. S. DeOrtentiis, P., K. Summers, J., P. Ammeter, A., Douglas, C., & R. Ferris, G. (2013). Cohesion and satisfaction as mediators of the team trust–team effectiveness relationship: An interdependence theory perspective. Career Development International, Vol. 18, No. 5, pp. 521-543.
4. Capretz, L. F. (2003). Personality types in software engineering. International Journal of Human-Computer Studies, 58(2), 207-214.
5. Capretz, L. F., & Ahmed, F. (2010). Making sense of software development and personality types. IT professional, 12(1).
6. Cockburn, A., & Highsmith, J. (2001). Agile software development, the people factor. Computer, 34(11), 131-133.
7. Graziotin, D., Wang, X., & Abrahamsson, P. (2014). Happy software developers solve problems better: psychological measurements in empirical software engineering. PeerJ, 2, e289.
8. Graziotin, D., Wang, X., & Abrahamsson, P. (2014). Software developers, moods, emotions, and performance. arXiv preprint arXiv:1405.4422.
9. Novielli, N., Calefato, F., & Lanubile, F. (2014, November). Towards discovering the role of emotions in stack overflow. In Proceedings of the 6th international workshop on social software engineering (pp. 33-36). ACM.
10. Ortu, M., Destefanis, G., Counsell, S., Swift, S., Tonelli, R., & Marchesi, M. (2017). How diverse is your team? Investigating gender and nationality diversity in GitHub teams. Journal of Software Engineering Research and Development, 5(1), 9.
11. Hertzum, M. (2002). The importance of trust in software engineers' assessment and choice of information sources. Information and Organization, 12(1), 1-18.
12. Moe, N. B., & Šmite, D. (2008). Understanding a lack of trust in Global Software Teams: a multiple□case study. Software Process: Improvement and Practice, 13(3), 217-231.
13. Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of

organizational trust. Academy of management review, 20(3), 709-734.
14. Deutsch, M. (1958). Trust and suspicion. Journal of conflict resolution, 2(4), 265-279.
15. Lewis, J. D., & Weigert, A. (1985). Trust as a social reality. Social forces, 63(4), 967-985.
16. Grabowski, M., & Roberts, K. H. (1999). Risk mitigation in virtual organizations. Organization Science, 10(6), 704-721.
17. Kanawattanachai, P., & Yoo, Y. (2002). Dynamic nature of trust in virtual teams. The Journal of Strategic Information Systems, 11(3-4), 187-213.
18. Martins, L. L., Gilson, L. L., & Maynard, M. T. (2004). Virtual teams: What do we know and where do we go from here?. Journal of management, 30(6), 805-835.
19. Jarvenpaa, S. L., Shaw, T. R., & Staples, D. S. (2004). Toward contextualized theories of trust: The role of trust in global virtual teams. Information systems research, 15(3), 250-267.
20. Ebert, C., Dumke, R., Bundschuh, M., & Schmietendorf, A. (2005). Best Practices in Software Measurement: How to use metrics to improve project and process performance. Springer Science & Business Media.
21. Rothenberger, M. A., Kao, Y. C., & Van Wassenhove, L. N. (2010). Total quality in software development: An empirical study of quality drivers and benefits in Indian software projects. Information & Management, 47(7-8), 372-379.
22. Gopal, A., Krishnan, M. S., Mukhopadhyay, T., & Goldenson, D. R. (2002). Measurement programs in software development: determinants of success. IEEE Transactions on software engineering, 28(9), 863-875.
23. Jeffrey, D. R., & Vessey, I. (1980). Models, metrics, and management of IS development. Information & Management, 3(3), 89-93.
24. 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.
25. Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. International journal of social research methodology, 8(1), 19-32.
26. Kitchenham, B. A., & Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Tech. rep., Technical report, EBSE Technical Report EBSE-2007-01.
27. C. B. Seaman, "Qualitative Methods". In: SHULL et al. (eds.), Guide to Advanced Empirical Software Engineering, Chapter 2, Springer, 2008.
28. Wohlin, C. (2014, May). Guidelines for snowballing in systematic literature studies and a replication in software engineering. In Proceedings of the 18th ACM international conference on evaluation and assessment in software engineering (p. 38).

Medición y Evaluación de Calidad de Sistemas Software para Flotas Dinámicas en Ciudades Inteligentes

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

(*) Universidad Rey Juan Carlos
(*) Grupo de Inteligencia Artificial (GIA). Centro para las Tecnologías Inteligentes de la Información y sus Aplicaciones (CETINIA) – Madrid – España
e-mail: alberto.fernandez@urjc.es

RESUMEN

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 [1].

Los nuevos sistemas de uso compartido de vehículos, llamados flotas dinámicas, son parte de este desafío que plantean las Ciudades Inteligentes [2]. 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.

Por tanto, la obtención de software 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.

Se muestra en este trabajo, el avance sobre lo que se ha estado trabajando: el desarrollo de una herramienta que permite automatizar la medición y evaluación de la calidad de aplicaciones web-mobile para este tipo de flotas, basada en un Modelo de Calidad cuyo pilar es la norma de calidad de producto ISO 25000 [3].

Palabras clave: Flotas dinámicas, Medición y Evaluación de calidad software.

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 Nº 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.

1. INTRODUCCIÓN

El software se ha incrustado profundamente en casi todos los aspectos de nuestras vidas y, como consecuencia, el número de personas que tienen interés en las características y funciones que brinda una aplicación específica ha crecido en forma notable [4].

Adicionalmente, la idea de ciudad inteligente ha presentado nuevos desafíos y requerido nuevas soluciones relacionadas al tráfico y el transporte. Como resultado directo, en los últimos años han surgido este tipo de sistemas que promueven el uso

compartido de vehículos [5]. Los servicios públicos de bicicletas, los sistemas compartidos de autos o bicicletas, o aplicaciones como UBER, proveyendo un servicio de taxi a través de conductores "libres", tienen el objetivo de mejorar la movilidad humana [6].

El servicio público de bicicletas, por ejemplo, ya está implementado en grandes ciudades alrededor del mundo como París, Barcelona, Londres, Shanghai, Nueva York, entre otras [7]. Una de las plataformas de bicicletas compartidas más grandes del mundo, OFO [8], 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.

Por esta razón, resulta esencial obtener un producto de alta calidad para la movilidad de los ciudadanos. Al agregar valor para el productor y para el usuario de un producto, el software de alta calidad proporciona beneficios a la organización que lo produce y a la comunidad de usuarios finales. Y como contraparte, el software de mala calidad aumenta los riesgos tanto para el desarrollador como para el usuario final [4].

La medición permite ganar comprensión acerca de un proceso y proyecto [4]. Por tanto, se necesitan herramientas eficientes que utilicen métodos efectivos para concretarla. Entonces, para la continuación de la construcción del módulo software en este trabajo se ha elegido la estrategia de Medición y Evaluación (M&E) GOCAME (Goal-Oriented Context-Aware Measurement and Evaluation) [9], con enfoque top-down u orientado a objetivos.

Particularmente, el enfoque top-down en esta estrategia se sigue de la siguiente manera: en primer lugar, se define el propósito de la M&E, a partir de éste se instancia un modelo de calidad adecuado, luego se identifican las métricas e indicadores más apropiadas que se utilizarán en la medición y la evaluación, y finalmente se analiza la información obtenida teniendo en cuenta el propósito inicial de la necesidad de información.

El proceso de M&E integra las siguientes seis actividades principales:

(A1) Definir los Requerimientos No Funcionales
(A2) Diseñar la Medición
(A3) Implementar la Medición
(A4) Diseñar la Evaluación
(A5) Implementar la Evaluación
(A6) Analizar y Recomendar

Cabe mencionar que GOCAME, además de seguir un enfoque orientado a objetivos, también tiene en cuenta el contexto en el cual se produce la Medición y Evaluación, con el fin de proporcionar información más precisa para la toma de decisiones, que es lo que se busca al medir la calidad de un producto.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En base a lo explicado anteriormente, en esta línea de investigación se viene trabajando en el desarrollo de una herramienta software que permite la recolección de datos desde una aplicación de bicicletas compartidas llamada Ecobike [10], desarrollada en un trabajo de fin de grado de alumnos en Ingeniería Informática de la Universidad Rey Juan Carlos, Madrid.

Siguiendo la estrategia de M&E elegida, a partir de las necesidades y objetivos de la aplicación, la primera actividad consiste en Definir los Requerimientos No Funcionales, lo cual incluye especificar el propósito, el foco y los atributos de la evaluación, entre otros aspectos. En este caso, la entidad concreta de estudio sobre la que se está trabajando es la aplicación de bicicletas compartidas, Ecobike [10], y el Modelo de Calidad seleccionado es un modelo específicamente desarrollado para la M&E de esta aplicación, un modelo mixto basado en la norma ISO 25000 [3].

El Diseño de la Medición, es decir, la identificación de las métricas que se utilizan para cuantificar los atributos, también se toman del modelo definido. Sólo se asigna una métrica para cada atributo del árbol de requerimientos [11].

Para la Implementación de la Medición, la herramienta provee la interfaz adecuada que permite el ingreso de los datos para realizar el cálculo de los valores finales.

Es necesario el Diseño de la Evaluación porque se necesita definir una transformación a partir del valor medido que produzca un nuevo valor numérico o categórico. Esta transformación está daba por un Indicador Elemental [12]. Asimismo, existen Indicadores Derivados, los cuales permiten interpretar los requerimientos de mayor nivel de abstracción, es decir, los conceptos calculables.

Bajo estas consideraciones, se define, para cada atributo y concepto calculable del árbol de requerimientos, un indicador que lo evaluará. Es decir, se define cómo se van a interpretar los valores de los atributos, para luego obtener el grado de satisfacción brindado por cada uno de los requerimientos de mayor nivel de abstracción.

Al definir un indicador derivado se debe indicar el modelo global (función o algoritmo) mediante el cual se calculará el nivel de satisfacción de los diferentes conceptos calculables del modelo de calidad.

Para el caso de estudio de Ecobike [10], se optó por utilizar el mismo modelo global para todos los indicadores. Particularmente, se escogió un modelo cuantitativo basado en el método LSP [13].

El método LSP se basa en la siguiente fórmula o función:

$ID(r) = (P1 * I1r + P2 * I2 r + ... + Pm * Im r)1/ r$

donde:

ID representa el valor del Indicador Derivado a calcular

Ii son los valores de los Indicadores del nivel inferior,

$0 <= Ii <= 100$

Pi representan pesos o ponderación,

$(P1 + P2 + ... + Pm) = 1; Pi > 0; i = 1 ... m$

r es un coeficiente conjuntivo/disyuntivo,

$-\infty <= r <= +\infty$

De acuerdo a esta función, se observa que existen dos parámetros a ser fijados: los pesos (Pi) y el coeficiente del bloque de agregación (r).

Para los pesos, la herramienta software permite asignar diferentes valores de importancia a cada atributo dentro de cada subconcepto, y, del mismo modo, a cada subconcepto que conforman el concepto calculable Calidad de Producto.

Los distintos valores de r, dependiente del tipo de operador y la cantidad de entradas, ya se encuentran tabulados en el modelo LSP [13].

En el caso de estudio, a nivel de subconceptos, el operador lógico empleado es C-, un operador para modelar una relación de cuasi-conjunción débil.

Al Implementar la Evaluación se obtienen los valores de los distintos indicadores asociados a cada uno de los elementos del árbol de requerimientos. Se puede conocer el grado de satisfacción proporcionado por los diferentes conceptos calculables y atributos de Ecobike.

Finalmente, se realiza la actividad Analizar y Recomendar. Teniendo en cuenta la necesidad de información o el problema que se desea atender, esta actividad permite realizar un análisis de las medidas y valores de indicadores obtenidos y recomendar acciones a seguir.

La herramienta en desarrollo está definida por los requerimientos que se detallan a continuación:

Medir la Calidad de Producto de una aplicación de movilidad para Smart City a partir del Modelo de Calidad.

- Calcular las métricas de cada atributo.
- Permitir detallar el peso de cada atributo.
- Calcular los indicadores elementales.
- Calcular los indicadores derivados.
- Interpretar los datos y generar reportes.

En base al análisis de estos requerimientos, se construyeron el diagrama de clases y el de casos de usos para su implementación posterior.

Con base en las actividades Diseñar la Medición y Diseñar la Evaluación, se definieron los elementos de interfaz para cada métrica e indicador.

Cada atributo del árbol de requerimientos definido por el Modelo de Calidad, queda

determinado por una pestaña que contiene los elementos de interfaz para cargar los datos que requieren las métricas. El cálculo de las mismas, de Indicadores Elementales y Derivados es automático.

Al finalizar con el cálculo de todas las métricas e indicadores elementales, se calculan los indicadores derivador y el indicador de calidad del producto y también se genera el reporte de análisis con los resultados obtenidos.

La aplicación Ecobike 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 de algunas de las métricas (por ejemplo: Tiempo de respuesta, para medir el Comportamiento Temporal), sin necesidad de tomarse el tiempo para realizar la observación y registro manual de los mismos.

Como resultado del proceso realizado mediante esta herramienta, se obtienen valores relevantes de análisis de la aplicación, 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.

3. RESULTADOS Y OBJETIVOS

La medición puede aplicarse al proceso de software con la intención de mejorarlo de manera continua, puede usarse a través de un proyecto de software para auxiliar en estimación, control de calidad, valoración de productividad y control de proyecto. Finalmente, la medición puede ser usada por los ingenieros del software para ayudar en la valoración de la calidad de los productos de trabajo y auxiliar en la toma de decisiones tácticas conforme avanza un proyecto [4].

Basados en estas declaraciones, se concluye que resulta indispensable la medición de la calidad de aplicaciones con tanto impacto en la gestión de la movilidad y tráfico de una ciudad, y la utilización de métodos/herramientas que permitan hacer un trabajo eficiente y práctico.

La aplicación del método propuesto en este trabajo a cualquier software de movilidad mediante una herramienta automatizada permite alinearnos con la meta dominante de la ingeniería del software: producir un sistema, aplicación o producto de alta calidad dentro de un marco temporal que satisfaga una necesidad de mercado. Para lograr esta meta, deben aplicarse métodos efectivos acoplados con herramientas modernas dentro del contexto de un proceso de software maduro [4].

El objetivo general sigue siendo 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, está el desarrollo del módulo software para la recolección de datos requeridos por las métricas e indicadores, y el cálculo de los mismos, y, también, diseñar y concretar la evaluación de calidad para poder introducir mejoras y optimizar el producto que se analiza.

Como contribución, se puede mencionar que la herramienta en desarrollo permite aportar a la optimización de la gestión del tráfico de una ciudad inteligente, debido a que posibilita la medición y evaluación de aplicaciones de movilidad, de forma tal que el uso de productos software de este tipo 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.

4. 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.

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.

Como así también se están llevando a cabo algunas tesinas de grado para la Licenciatura en Computación. Y trabajo finales de carrera de la Ingeniería en Informática e Ingeniería en Computación de la Facultad de Ciencias Físico-Matemáticas y Naturales, Universidad Nacional de San Luis.

5. REFERENCIAS

[1] G. Graham and L. Zhang, "Smart cities and digital technologies: the case of bike-sharing systems," *http://eprints.whiterose.ac.uk/84589/1/smart%20cities%20and%20bike-sharing%202015-1.pdf* 2015.

[2] O. Mont, "Institutionalisation of sustainable consumption patterns based on shared use.," *Ecological economics, The International Institute for Industrial Environmental Economics at Lund University.,* pp. 135- 153, 2004.

[3] ISO/IEC 25010:2011. Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models. ."

[4] R. S. Pressman., *Ingeniería del Software: un enfoque práctico. Séptima Edición*, 2006.

[5] O. Mont. Institutionalization of sustainable consumption patterns based on shared use. Ecological economics, 50(1):135–153, 2004

[6] Holger Billhardt, Alberto Fernández, Marin Lujak, Sascha Ossowski, Vicente Julian, Juan F. De Paz, Josefa Z. Hernandez. "Coordinating Open Fleets. A Taxi Assignment Example", 2016

[7] Kabra, Ashish and Belavina, Elena and Girotra, Karan, Bike-Share Systems: Accessibility and Availability, 2015. Chicago Booth Research Paper No. 15-04.

[8] www.ofo.com/us/en

[9] Olsina, L., Papa, M., & Molina, H. How to Measure and Evaluate Web Applications in a Consistent Way. In Rossi, Pastor, Schwabe, & Olsina (Eds.), Web Engineering: Modeling and Implementing Web Applications (pp. 385– 420). Springer HCIS. 2008.

[10] Ecobike Solutions: Gómez Pérez, Manuel y López González, Julio. Gestión de sistemas de alquiler de bicicletas. 2016.

[11] Covella, Guillermo, Dieser, Alexander y Olsina, Luis. Una Estrategia de Medición y Evaluación como Soporte para la Gestión de Tecnologías de Información en el Estado. 2013.

[12] Olsina, L., Martin, M. Ontology for Software Metrics and Indicators, Journal of Web Engineering , RintonPress, US, 2004.

[13] J. J. Dujmovic, "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.

Metamodelo Ontológico como Garantía de Calidad en el Desarrollo de Software

Rebeca Yuan*, Carlos Salgado+, Alberto Sánchez+, Mario Peralta+
*Ingeniería de Software, Departamento en Ingeniería en Sistemas de Información
Facultad Regional San Francisco – Universidad Tecnológica Nacional
ryuan@sanfrancisco.utn.edu.ar
+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

La calidad en el desarrollo del software, es un elemento "distintivo" para aquellas organizaciones que buscan lograr posición y competitividad en el mercado [1, 2]. La implementación de la misma, implica buenas prácticas, desde la extracción de requerimientos hasta su testeo y entrega; siendo la primera etapa del proceso la base fundamental para asegurar el éxito del software desarrollado [2,3]. La extracción de requerimientos del software no es tarea sencilla; la interpretación del cliente, la transferencia hacia los desarrolladores, y su evaluación final, genera la necesidad de distintas herramientas que den soporte a éstas tareas y colaboren con la gestión de calidad. El presente trabajo busca analizar y definir cómo el uso de Modelos Ontológicos logra mejorar la etapa de educción de requerimientos del software, colaborando con las distintas normas de calidad que las empresas establecen; garantizando que en la etapa de testeo los errores conceptuales disminuyan.

Palabras Claves: *Calidad de Software, Modelos Ontológicos, Educción de Requerimientos, Machine Learning.*

CONTEXTO

El presente es un trabajo realizado en conjunto entre el Grupo de Investigación sobre Aplicaciones Inteligentes (GISAI), de la Universidad Tecnológica Nacional Facultad Regional San Francisco bajo Resolución de CD N°1126/2015 de fecha 9/11/2015 y 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 PROICO Nº P-031516.

1.INTRODUCCIÓN

La velocidad en el procesamiento de la información otorga al desarrollo del software un marco potencial para la creación de nuevos sistemas. Organizaciones y empresas que brindan servicios, buscan incorporar nuevas herramientas tecnológicas que otorguen un valor agregado a sus clientes. Esta posibilidad de contar con hardware preparado para soportar grandes procesamientos, abrió las puertas a la Inteligencia Artificial, y con ella a herramientas basadas en aprendizaje automático (Machine Learning)[4], que avanzaron un paso más en las posibilidades de servicio que puede ofrecer un negocio. Conocer las necesidades de los clientes, sus gustos y tendencias, mejorar la comunicación (a través de *chatbot)*, optimizar la seguridad en la información de las empresas, automatización y desarrollo de servicios; podemos nombrar infinidad de beneficios, pero nunca podemos olvidar, que el inicio de

cualquiera de las soluciones nombradas anteriormente, parten de una necesidad por parte de nuestros clientes, *sus requerimientos*. Según una de las especificaciones de la IEEE un requerimiento es: "Una condición o capacidad que debe estar presente en un sistema o componentes de sistema para satisfacer un contrato, estándar, especificación u otro documento formal" (Std 610,12-1900,IEEE:62)[5]; el cumplimiento de esta definición es base para verificar si se cumplen o no las necesidades del o los clientes, siguiendo un estándar de calidad. El descubrir, saber interpretar e implementar esa solicitud es lo que distingue a un desarrollo exitoso. Algunas organizaciones tienen la posibilidad de encontrarse acompañadas por profesionales expertos en la extracción, definición y análisis de requerimientos que no solo descubren lo que el cliente busca, sino que saben proyectar a estas empresas para tiempos futuros, explotando esos requerimientos, para contar el día de mañana con información que alimente sistemas inteligentes [6]. La complejidad del problema a resolver, la comunicación entre clientes, usuarios y desarrolladores, el descubrimiento de requerimientos ocultos, son algunos de los problemas que se enfrentan a la hora de comenzar el ciclo de vida del software.

La implementación de normas de calidad en el desarrollo de software aporta buenas prácticas de trabajo y herramientas que establecen un orden y organización para la captura de datos. La familia de normas ISO/IEC 25000, conocida como SQuaRE (*System and Software Quality Requirements and Evaluation*), ofrece un marco de trabajo común para evaluar la calidad del software. El Modelo de Madurez y Capacidad Integrado (CMMI), también brinda buenas prácticas que van desde el desarrollo y mantenimiento, hasta la puesta en marcha del producto de software [1,2,3]. Si bien los beneficios que brinda la implementación de Normas de Calidad es importante, no todas las empresas de software acceden a dicha implementación. Por este motivo, y siguiendo los lineamientos propuestos por las normas de calidad mencionadas, la incorporación de *ontologías* en la captura de requerimientos y modelado de negocio lograría mejorar el entendimiento compartido de los actores del proyecto (cliente-analista-desarrollador), reduciendo el número de errores conceptuales y tiempos en la etapa de desarrollo, permitiendo la trazabilidad de cada requerimiento.

La ontología es una disciplina muy antigua; que es aplicada en distintos ámbitos. En la antigüedad, ésta significaba la comprensión general del ser, esto dentro de una perspectiva metafísica. Buscando alejarnos de esta mirada inicial y filosófica, Gruber en 1993 [7,8] la establece como "Una especificación formal, explícita de una conceptualización compartida", por lo expresado esta disciplina aporta un lenguaje común, reduciendo la ambigüedad de conceptos, logrando un entendimiento compartido sobre distintos problemas, además de facilitar la comunicación entre los actores del proyecto [9]. Este modelo semántico ofrece una abstracción sobre un área de conocimiento. Estas cualidades positivas hacen a la ontología una herramienta importante dentro del proyecto de software, muchas veces ignorada por la Ingeniería del Software. La ontología también forma parte de una rama de la Inteligencia Artificial, se la encuadra también dentro de la Ingeniería del Conocimiento [6]. Existen muchos antecedentes de aplicación de ontología en la extracción de requerimientos y muchos investigadores se encuentran buscando una ontología genérica o meta-ontología que permita una extracción óptima de los mismos, yendo desde lo general hasta el detalle de los requisitos. De acuerdo al nivel de generalidad, Guarino [10] propone una clasificación de Ontologías de Alto Nivel, (así

también otras de Dominio, de Tareas y de Aplicación); la cual permite describir conceptos generales independientes del dominio o problema; en una investigación realizada (Aranda & Ruiz, 2005) [11,12] se clasifica y comparan distintas ontologías de diferentes autores; la investigación que nos compromete busca comenzar por una ontología de alto nivel de generalidad para luego bajar a conceptos del proceso o tarea; una ontología de orden general (Meta-ontología), compuesta por distintas categorías que a su vez se dividen en niveles más bajos.

El uso de las ontologías no solo mejora el entendimiento entre los distintos actores que forman parte del proyecto de software, sino que también persigue a través de esta ontología genérica la usabilidad de la misma en distintos dominios.

En base a lo expuesto, en la presente línea de investigación se está trabajando en el Análisis y definición de Ontologías y el establecimiento de una Meta-ontología para la Educción de Requerimientos.

2.LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

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

- Ingeniería de Software.
- Calidad de Software.
- Sistemas Inteligentes.

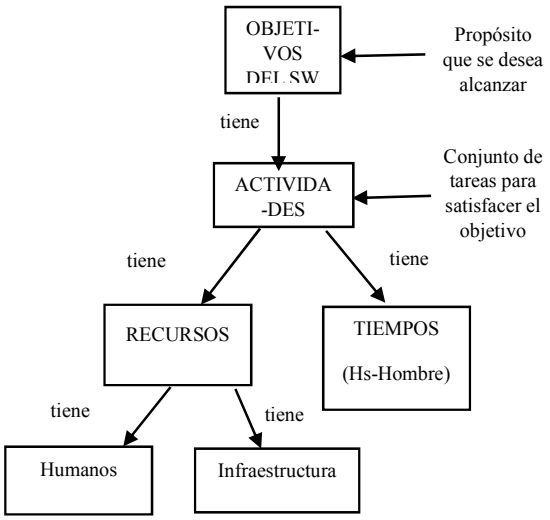
El presente proyecto de investigación cuenta con varias etapas: 1) Estado del arte de la Ontología, 2) Análisis de los distintos artefactos que surgen de la aplicación de normas de calidad (ISO/ CMMI) para la extracción de requerimientos; 4) Estudio de Meta-Ontologías para la extracción de Requerimientos. 5) Aplicación de Modelos Ontológicos para la Educción de Requerimientos 6) Validación de la propuesta ontológica.

3.RESULTADOS ESPERADOS

En paralelo a la primera etapa definida como parte del trabajo de tesis (Estado del Arte de la Ontología), se realizó un relevamiento a cinco (5) empresas de desarrollo de software de la ciudad y región. La consulta estuvo dirigida a Analistas, Desarrolladores y encargados de Testing; de estos actores, quien tiene la primera entrevista con el cliente o con el área que requiere el software, es el Analista (o Ing. En Sistemas). En todas las empresas consultadas se cuenta con personal experto y con más de seis años de antigüedad en el puesto; en ésta etapa se identifica el objetivo general del desarrollo, los tiempos aproximados de trabajo y los recursos que se necesitan para ejecutarlos. Continuando, el cien por ciento (100%) trabaja la extracción de requerimientos en *texto plano* y sólo un 5% intenta aplica metodologías ágiles para dar soporte a la extracción y seguimiento de los mismos. La mayoría pretende seguir algún estándar de calidad o buenas prácticas de desarrollo (pero no buscan la certificación); emplean herramientas propuestas por metodologías como Scrum o estándares propuestos por IEEE (Historias de usuario, Documentos estáticos, etc). Una de las empresas consultadas, trabaja en la parte de testing con la certificación internacional otorgada por ISTQB (International Software Testing Qualifications Board) pero al momento de implementarla se topan con un problema en la etapa de extracción de requerimientos; los mismos se encuentran poco documentados o explícitos, generando así problemas para determinar si lo que el cliente pidió fue lo desarrollado en forma correcta; esta forma de trabajo no vislumbra la posible aplicación de una solución inteligente a futuro.

Queda así determinado; qué para todas las empresas entrevistadas, todos los proyectos de software comienzan con la "entrevista" al cliente; es la principal fuente de extracción de información, continuando con el análisis de documentos propios de cada negocio, toda la información recolectada queda detallada en forma textual por los analistas, para ser enviada a las siguientes etapas del ciclo del proyecto.

Con la información adquirida, podemos plantear una primera versión de Meta-Ontología y verificar la propuesta por Guarino donde establece que "se describen conceptos generales como espacio, tiempo, acción…"[13]. Los componentes de una ontología también varían según distintos autores, pero en grandes rasgos, se presentan como axiomas, clases o tipos, instancias o individuos, relaciones, propiedades, etc.



Primera versión de Meta-ontología propuesta

El cuadro muestra la primera propuesta de Meta-ontología, obtenida de las encuestas realizadas; queda pendiente la validación de la misma y su mejora.

El trabajo continuará con el planteo de distintas herramientas ontológicas que cubran la lingüística utilizada en las distintas entrevistas con los clientes; como ser el estudio del Universo del Discurso (UdeD) y el Léxico Extendido del Lenguaje (LEL)[12]; buscando identificar todo el vocabulario y su semántica; práctica recomendada por la IEEE para las especificaciones de los requisitos de software y distintas herramientas que enriquezcan el entendimiento del lenguaje. Se cerrará la investigación con la validación de la ontología establecida.

4.FORMACIÓN DE RECURSOS HUMANOS

En el marco de esta línea de investigación, se está desarrollando una tesis para la Maestría en Calidad de Software dictada por la Universidad de San Luis bajo Resolución ME N° 1589/13. En el ámbito de dicho proyecto se propone la investigación sobre el impacto de un Meta-modelo Ontológico para la extracción de requerimientos del Software. Como docente de la Cátedra Inteligencia Artificial, la implementación de la ontología establece un marco necesario para el análisis y estructura de datos; base necesaria para el desarrollo de herramientas basadas en Machine Learning (Aprendizaje automático).

En forma paralela, se está trabajando en el Monitoreo de Carga No Intrusiva (NILM) implementando una Red Neuronal Artificial (RNA) para la identificación de artefactos eléctricos. La aplicación de herramientas de Machine Learning en distintos tipos de dominio es uno de los objetivos del grupo de investigación. El equipo 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.

Como iniciativa de este trabajo, y con el objeto de volcar lo aprendido en el 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 (Ontologías, Ciencia de Datos, M*achine 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, a la realización de actividades de investigación y desarrollo.

5.BIBLIOGRAFÍA

[1] Piattini Velthuis, M. G., García Rubio, F. O., García Rodriguez de Guzmán, I., & Pino, F. (2011). Calidad de Sistemas de Información. México: Alfaomega Ra-Ma.

[2] Sangüesa, M., Mateo, R., & Hzarbe, L. (2006). Teoría y Práctica de la Calidad. Madrid, España: Thomson.

[3] Ishikawa, K. (1994). Introducción al Control de Calidad. Madrid, España: Díaz de Santos S.A.

[4] Mitchell, T. M. (1997). Machine Learning. McGraw - Hill Science.

[5] Electrónica, I. d. (s.f.). Std 610,12-1900,IEEE:62. Obtenido de https://www.ieee.org/

[6] Russell, S., & Norvig, P. (2010). Artificial Intelligence A Modern Approach. Prentice Hall Series - Pearson.

[7] Gruber, T. R. (1993). A translation approach to portable ontology specification.

[8] Arias, C. M. (2005). La ingeniería de requerimientos y su importancia en el desarrollo de proyectos de software. InterSedes@Universidad de Costa Rica, Vol. 6.

[9] Zapata , C. M., Giraldo, G. L., & Urrego Giraldo, G. A. (2010). Las ontologías en la ingeniería del software: un acercamiento de dos grandes áreas de conocimiento. Ingenierías Universidad de Medellín, 91-99.

[10] N., Guarino. (1998). Formal Ontology in Information Systems. Amsterdam: IOS Press.

[11] Aranda, G. N., & Ruiz, F. (2005). Clasificación y ejemplos del uso de ontologías en Ingeniería del Software.

[12] Zapata, C. M., Giraldo, G. L., & Mesa, J. E. (2010). Una propuesta metodológica para la educción de requerimientos. INGENIARI, pp. 26-37.

[13]García-Peñalvo, F. J. (s.f.). Web Semántica y Ontologías. Obtenido de ResearchGate: https://www.researchgate.net/publication/267222548_Web_Semantica_y_Ontologias.

Métodos y Técnicas de Evaluación de Interfaz Gráfica de Usuario

M. Claudia Albornoz, Mario M. Berón, Germán Montejano
Departamento de Informática/Universidad Nacional de San Luis-U.N.S.L./San Luis/Argentina
Ejército de los Andes 950, Tel: +54 (0266) 4520300; int 2102
{albornoz,mberon,gmonte}@unsl.edu.ar

Resumen

El mundo actual se mueve y comunica a través de la tecnología. Es fundamental que las personas puedan interactuar con los nuevos productos tecnológicos, sistemas de software, sitios de internet. Por lo antes mencionado, lograr que el usuario inexperto pueda comunicarse de manera fluida con los productos tecnológicos es un desafío relevante. Como punto de partida, se debe mostrar al usuario una Interfaz accesible, amigable, satisfactoria; que le permita concretar trámites y tareas fácilmente, lo cual no es una tarea menor.

El presente trabajo describe la línea de investigación que aborda el estudio de métodos y técnicas utilizados para evaluar Interfaces Gráficas de Usuario.

Palabras clave: Interfaz Gráfica de Usuario, Evaluación, Métodos, Técnicas.

Contexto

El presente trabajo se enmarca en el Proyecto de Investigación: "Ingeniería de Software: conceptos, prácticas y herramientas para el desarrollo de software de calidad" código P-031516 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.

Introducción

En los últimos tiempos, todas las personas nos vemos en la necesidad, y casi obligación, de interactuar con la tecnología. Los productos tecnológicos forman parte de la vida cotidiana de las personas. No solo en la actividad profesional o laboral, sino también en el hogar o para concretar tareas simples como extraer dinero de un cajero, comunicarse con otras personas (llamadas, mensajes, redes sociales, etc.), solicitar un turno en instituciones gubernamentales o sistemas de salud pública, etc. Así la lista puede enumerar una gran cantidad de tareas. Para el usuario inexperto, sin mayores conocimientos, estas tareas pueden ser complicadas y engorrosas si las Interfaces con las que interactúan no facilitan la tarea.

Se puede definir Interfaz como: Conjunto de formas y métodos que posibilitan la interacción de un sistema con los usuarios utilizando formas gráficas e imágenes. Con formas gráficas se refiere a botones, íconos, ventanas, fuentes, etc. los cuales representan funciones, acciones e información [1].

Se espera que el usuario se encuentre con una Interfaz accesible, amigable, eficiente, efectiva, satisfactoria, etc. Esto se logra si la Interfaz posee un 'buen diseño', si cuenta con una característica fundamental que es la

‘Usabilidad’; es decir si cuenta con aquella característica ‘que hace que la aplicación sea fácil de utilizar y fácil de aprender’ [2].

Las características antes mencionadas y otras (como: robustez, consistencia, ayuda al usuario, organización, claridad, etc.) forman parte de una extensa lista denominadas ‘Características Deseables’. Establecer ésta lista de características tiene un objetivo muy importante, el cual es poder evaluar las Interfaces Gráficas de Usuario.

La evaluación sirve para mejorar, descubrir errores o fallas y buscar la manera de solucionarlos. La evaluación ayuda a ver las equivocaciones, y al detectarlas se puede ser consciente de los errores y esto ayuda a aprender.

El conocimiento que se adquiere en el proceso de evaluación seguido de la corrección de fallas es lo que ayuda al programador a mejorar el diseño del producto. Esto lleva a una retroalimentación dado que se puede diseñar mejor si se evalúa, se detectan errores y se aprende de ellos [3].

La evaluación es útil en términos de:

- Mejoras en cuestiones de diseño.
- Identificación de problemas.
- Mejoras en la calidad.
- Ayuda a seleccionar la mejor opción en caso de disponer de diversas aplicaciones para realizar una tarea.

Éste último punto es una de los principales objetivos de la evaluación; al contar con diferentes productos para realizar una misma tarea, la evaluación de la Interfaz es quien puede determinar cuál producto es el más apto para trabajar. Aquella que se muestre más accesible, entendible, que ofrezca una interacción fluida con el usuario entre otras opciones, sería la elegida para realizar la tarea [4].

A continuación, se mencionan y explican algunos métodos y técnicas de evaluación.

1.- Métodos y Técnicas

1.1- Métodos de Inspección

a) Evaluación Heurística: partiendo del término ‘heurística’, el diccionario de la RAE lo define como:

- una técnica de indagación y descubrimiento.
- una manera de buscar la solución a un problema mediante métodos no rigurosos.

Esta técnica de evaluación, aplicado a las Interfaces debe ser llevada a cabo por evaluadores expertos a partir de principios llamados heurísticos establecidos con anticipación. Es lógico pensar, que los evaluadores pondrán mayor énfasis en lo que suponen percibirán los usuarios.

La evaluación heurística inspecciona problemas potenciales, ya que el evaluador predice los errores que podría tener el usuario común [3].

El éxito o fracaso de éste tipo de evaluación depende en gran medida de las heurísticas establecidas para hacer la evaluación. Una vez establecidas, los expertos individualmente, examinan si la Interfaz evaluada cumple con cada heurística, asignándole un puntaje. Luego se comparan los puntajes y se confecciona un informe [5]. Este informe describe los problemas encontrados, ayuda al equipo de desarrollo a revisar los problemas y buscar una solución, con el objetivo de mejorar el diseño.

b)Recorrido cognitivo (cognitive walkthrough) : el objetivo central de ésta

técnica es evaluar la facilidad de aprendizaje a través de prototipos del sistema; puede aplicarse desde las etapas iniciales del desarrollo. Es importante que al aplicar ésta técnica se cumplan las diferentes etapas:

* Datos necesarios para el recorrido:
  - Identificar, documentar características del usuario final (cuánta experiencia y conocimiento poseen son factores determinantes para la comprobación del factor "cognitivo" durante el recorrido).
  - Describir el prototipo del sistema.
  - Enumerar las tareas a realizar.
  - Listar las acciones a realizar en cada tarea.

* Recorrer las acciones: cada evaluador debe realizar las tareas que fueron enumeradas antes, en este proceso el evaluador utilizará la información del factor cognitivo (conocimiento y experiencia adquirida) de los usuarios para comprobar si la interfaz es adecuada para los mismos. Esta revisión ha de ser minuciosa para todas las acciones especificadas como necesarias para la consecución de la tarea.

* Documentar resultados:
  - Cada evaluador documentará para cada acción las respuestas del sistema
  - El documento incluirá un anexo especial, conocido como "Usability Problem Report Sheet" en donde se detallen los aspectos negativos de la evaluación, asignándole un grado de severidad que permita distinguir entre errores graves de aquellos más leves [6], [7].

1.2 - Métodos de Indagación

Este tipo de métodos consiste en hablar con los usuarios y observarlos usando el sistema en trabajo real (se involucra directamente al usuario). Los principales técnicas de evaluación por indagación son:

a) *Entrevista:* usada para conocer la opinión de los usuarios. Son técnicas exploratorias y en ningún caso pueden constituir una medición formal. El principal aporte es que permite conocer el grado de satisfacción que tiene el usuario con el producto de software y sus valoraciones sobre los contenidos.
b) *Cuestionario:* técnica exploratoria de usos y motivaciones de los usuarios que permite conocer preferencias sobre contenidos, momentos de conexión, familiaridad con el producto e intereses. No es una técnica formal para medir.

1.3 - Métodos Multicriterios

Los métodos multicriterios se utilizan, como su nombre lo indica, en función a varios criterios; los cuales se pueden medir. La idea es determinar un conjunto de criterios o características a evaluar. Dada la necesidad de medir hasta dónde se satisfacen las características, es necesario disgregarlas una y otra vez en subcaracterísticas, y llegar a características atómicas.

Los métodos de evaluación multicriterios comprenden la selección entre un conjunto de alternativas factibles, la optimización con varias funciones objetivo simultáneas y un único agente decisor, y procedimientos de evaluación racionales y consistentes [8] [9]. Existen diferentes métodos multicriterios discretos; se pueden mencionar:

a) *Análisis Jerárquico o Proceso de Jerarquía Analítica (AHP: Analytic Hierarchy Process):* es un proceso lógico y estructurado, que optimiza la toma de decisiones complejas cuando

existen múltiples criterios, disgregando el problema en una estructura jerárquica. Esto permite subdividir un criterio complejo en un conjunto de atributos sencillos y determinar cómo influyen cada uno de esos atributos individuales en el objetivo de la decisión. Esa influencia está representada por la asignación de los valores que se otorga a cada atributo o criterio. El método AHP establece dichos valores a través de comparaciones por pares (uno a uno). En determinadas circunstancias esto facilita la objetividad del proceso y permite reducir sustancialmente el uso de la intuición en la toma de decisiones.

b) *Ponderación Lineal (scoring):* es el método más conocido y utilizado. Consisten en construir una función de valor para cada alternativa. Los métodos de Ponderación Lineal permiten obtener una puntuación global a partir de la suma de las contribuciones obtenidas de cada característica atómica (ésto se logra luego de un proceso de normalización para que puedan sumarse éstos valores). Son métodos compensatorios, pueden depender de la asignación de pesos a los criterios o de la escala de medida de las evaluaciones. Son métodos sencillos de utilizar, intuitivos y muy difundidos. LSP (Logic Scoring of Preference) es un método de evaluación multicriterio que se puede adaptar a diferentes contextos de evaluación (no solo sistemas de software) y permite la realización de evaluaciones precisas [10].

Hasta aquí se ha hecho una exposición de diferentes técnicas de evaluación. Dada la acotada extensión de los artículos, sólo se mencionan algunos de los tantos métodos y técnicas. El objetivo es mostrar la variedad de formas de evaluar Interfaces Gráficas de Usuario.

Líneas de Investigación, Desarrollo e Innovación

Los principales ejes de ésta línea de investigación son:

- Fundamentar la importancia de la evaluación de Interfaces Gráficas de Usuario, resaltando cuál es el objetivo de realizar dicha evaluación.
- Mencionar las diferentes metodologías de evaluación y las técnicas que incluyen las mismas: con la intervención de expertos, o con los usuarios, o con métodos Multicriterios, entre otras.
- Elaborar una lista de criterios fundamentales que deben cumplir las interfaces gráficas a fin de su construcción sea de calidad.
- Elaborar criterios de construcción y evaluación de interfaces gráficas específicas de domino. En esta línea de investigación se centra en la construcción de interfaces gráficas de personas con capacidades diferentes. Se trata de analizar la posibilidad de facilitarle el acceso y manejo de software a dichas personas como una forma de inclusión en la sociedad.

Resultados y Objetivos

La evaluación de las Interfaces Gráficas de Usuario no solo ayuda a seleccionar la mejor opción, o la más adecuada en cierta circunstancia cuando se debe utilizar una herramienta; sino también a mejorar el

diseño y la calidad, identificar problemas de diseño, de conceptos, de programación, etc. al momento de desarrollar un sistema.

Asiduamente se cree que la tarea del equipo desarrollador sólo se reduce a cumplir con los requerimientos del usuario, sin considerar características fundamentales que debe poseer una Interfaz; parte fundamental del producto, dado que es la parte visible y con quien interactúa y se comunica el usuario.

En este contexto se puede decir que se está desarrollando un reporte técnico donde se aborda la temática de esta línea de investigación en profundidad. A partir de este informe se ha elaborado una lista de criterios, los cuales están bajo revisión por profesionales informáticos, que son útiles para evaluar interfaces gráficas de usuario. Además de los resultados mencionados es importante destacar el desarrollo de dos herramientas que implementan métodos de evaluación multicriterio uno es NESSy [11] una herramienta que implementa el método LSP y la otra es DAST la cual implementa el método AHP [12]. Como se podrá observar la investigación se ha orientado a hacia los métodos de evaluación multicriterio, esto se deba a que el equipo de investigación percibe que se adaptan mejor al tipo de evaluación que se desea llevar adelante. No obstante, no se descarta el desarrollo de herramientas que no se encuentren en esta categoría, todo dependerá del rumbo que marque la investigación.

Formación 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, maestrías o especializaciones 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.

Referencias

[1] http://www.alegsa.com.ar/Dic/gui.php

[2] Galitz, W. O. *The essential guide to user interface design: an introduction to GUI design principles and techniques.* Wiley.com, 2007.

[3] Rodríguez, A. C. F., Santamaría, S. O., Saez, A. V., & Lluch, M. Z. (2011). *Evaluación de la usabilidad*. Universitat Oberta de Catalunya.

[4] Albornoz, M. C., Miranda, E., and Beron, M. (2013). *Evaluacion de interfaces gráficas de usuario usando LSP* - CoNaIISI 2013.

[5] DÍAZ, Fierro; YESENIA, Natali. *Heurísticas para evaluar la usabilidad de aplicaciones web bancarias*. 2015. Tesis para optar el grado de Magíster en Informática con mención en Ingeniería de Software.

[6] CANCIO, Lilliam Perurena; BERGUES, Mercedes Moráguez. *Usabilidad de los sitios Web, los métodos y las técnicas para la evaluación*. Revista Cubana de Información en Ciencias de la Salud (ACIMED), 2013, vol. 24, no 2, p. 176-194.

[7] GRANOLLERS, T.; PERDRIX, F.; LORÉS, J. *Incorporación de usuarios en la evaluación de la usabilidad por recorrido cognitivo*. Interacción'04, 2004.

[8] Martínez, E. y Escudey, M. *Evaluación y decisión multicriterio: reflexiones y experiencias*. Unesco, 1998.
[9] Albornoz, María Claudia; Berón, Mario; Montejano, Germán Antonio. *Evaluación de interfaces gráfica de usuario*. En XVII Workshop de Investigadores en Ciencias de la Computación (Salta, 2015). 2015.
[10] BERUMEN, Sergio A.; REDONDO, Francisco Llamazares. *La utilidad de los métodos de decisión multicriterio (como el AHP) en un entorno de competitividad creciente*. Cuadernos de administración, 2007, vol. 20, no 34.
[11] Miranda, E., Berón, M., Montejano, G., Pereira, M. J., & Henriques, P. (2013). *NESSy: a new evaluator for software development tools*. In *2nd Symposium on Languages, Applications and Technologies (SLATe'13)* (Vol. 29, pp. 21-38).
[12] Marini, Ayrthon; Miranda, Enrique; Berón, Mario; Bustos, Miguel; Riesco, Daniel; Henriques, Pedro. *Una Herramienta para la Evaluación de Sistemas de Software: Aplicación a la Evaluación de Escáneres de Seguridad en Aplicaciones Web.*
6° Congreso Nacional de Ingeniería Informática/Sistemas de Información (CoNaIISI 2018)

Métodos y Técnicas para Cálculo del Grado de Entendimiento de Servicios Web Basados en WSDL

Hernán Bernardis(1), Edgardo Bernardis(1), Mario M. Berón(1), Daniel E. Riesco(1), Pedro Rangel Henriques(2), Maria Joao V. Pereira(3)

(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/Universidade do Minho
Braga – Portugal
pedrorangelhenriques@gmail.com

(3) Departamento de Informática e Comunicações/ Instituto Politécnico de Bragança
Bragança - Portugal
mjoao@ipb.pt

Resumen

La transformación del desarrollo de software es cada vez mayor migrándose la mayoría de los sistemas a arquitecturas orientadas a servicios en la nube. Esta relativamente nueva modalidad de desarrollo de sistemas de software se hace difícil comprender los sistemas debido a que la complejidad de un sistema ahora radica en la complejidad de cada uno de sus módulos sumada a la complejidad de los servicios web con los que el mismo interactúa.

En el caso particular de los servicios web que utilizan un WSDL (Web Services Description Language) [1] para ser descriptos, la complejidad de entendimiento para un ingeniero de software se incrementa debido a que los mismos son diseñados para ser interpretados por agentes de software y no por personas. Esta alta formalidad permite incluso que estos servicios web posean herramientas de generación automática. Esta automatización resulta extremadamente útil y eficiente al reducir los tiempos de desarrollo pero carece de semántica que permita comprenderlo luego de creado.

En este artículo se describe una línea de investigación centrada el cálculo de múltiples métricas de los Servicios Web mediante el análisis de sus especificaciones WSDL. Estas métricas sirven como base para el cálculo del grado global de entendimiento del WSDL.

Palabras clave: Web Services, métricas, comprensión, WSDL.

CONTEXTO

La línea de investigación descripta en este artículo se desarrolla en el **La**boratorio de **C**alidad e **I**ngeniería de **S**oftware (**LaCIS**) de la Universidad Nacional de San Luis; y se encuentra enmarcada dentro del proyecto (P-031516): *"Ingeniería del Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de Calidad"*, perteneciente a la universidad antes mencionada. Dicho proyecto es reconocido por el programa de incentivos, y es la 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*", aprobado por el Ministerio de Ciencia Tecnología e Innovación Productiva (Mincyt) con código PO/16/93.

1. INTRODUCCIÓN

El incremento en el uso de los servicios web dentro del desarrollo de sistemas es cada vez mayor. Los sistemas migran hacia arquitecturas orientadas a servicios [2,3] en lugar de seguir con la arquitectura centralizada tradicional. La idea detrás de esto es similar a la modularización tradicional en el desarrollo de sistemas. En lugar de aislar código en clases, paquetes o métodos, para ser reutilizados mediante su llamada, en múltiples partes del sistema, se lo hace mediante la creación de servicios web. Esto posee dos grandes ventajas: i) su alcance es mucho más global debido a que los servicios web se publican en internet y son accesibles desde cualquier lado y ii) su construcción es independiente del lenguaje utilizado por quien lo invoque. Estas características permiten también que se puedan utilizar servicios web de terceros, es decir, sin necesidad que los mismos sean desarrollados por quienes lo utilizan.

La gran migración hacia arquitecturas orientadas a servicios ha generado múltiples herramientas y técnicas para automatizar la creación de las descripciones de los mismos. En el caso de los servicios web basados en WSDL, la formalidad de los WSDLs permite automatizar su construcción. Si un WSDL es creado con la formalidad necesaria para ser interpretado por un agente de software automáticamente, entonces suena lógico que pueda ser generado por otro software de manera automática.

Esto es altamente provechoso por el ahorro de tiempo que produce construir un sistema que realiza una determinada funcionalidad y, de manera automática, crear su interfaz y descripción para su uso como servicio web. Disminución de tiempos, evitarse la construcción de formalidades innecesarias, facilidad de construcción son algunas de las ventajas que proveen estas herramientas de generación automática. Pero, como todo, también hay desventajas en este proceso y las mismas aparecen a la hora de analizar los WSDLs generados.

Si se supone la situación particular de un sistema de una multinacional con múltiples invocaciones a servicios web que se debe migrar hacia una tecnología nueva. Este proceso lleva un planeamiento de migración en donde se deben especificar por cada módulo o funcionalidad los periodos en tiempo y los recursos necesarios para la migración. Para los sistemas tradicionales existen múltiples formas de estimar la complejidad [4,5] pero para los servicios web y los módulos afectados, que tan difícil será?, ¿Cuánto tiempo llevará?. Otra situación que puede ocurrir es el caso donde un banco posee múltiples servicios web en su sistema con arquitectura orientada a servicios y desea determinar qué tan vulnerables son a posibles ataques informáticos. ¿Cómo puede analizar las vulnerabilidades del mismo? ¿Cómo determinar que tan comprensibles son a personas externas que deseen encontrar vulnerabilidades en sus descripciones para atacar? Ambos escenarios desnudan la necesidad de técnicas de comprensión de servicios web, pilares funcionales de todo sistema.

Toda comprensión siempre nacerá con una extracción de información. Si no obtengo información de algo, ¿Cómo puedo llegar a comprenderlo? Es bien sabido que la comprensión dentro del proceso de desarrollo de software es la parte más costosa en tiempo y recursos, justamente por la dificultad que implica extraer la información adecuada y necesaria para realizarla [4,5]. Poca información no alcanzará, mucha información entorpecerá. Entonces saber qué información extraer es casi tan importante como la información en sí. En el caso particular de los

servicios web nace otro inconveniente: de dónde sacar la información de una funcionalidad encapsulada en algún lugar de internet que sólo nos dice su forma de usarla [6,7]. Es aquí donde la atención se desplaza, obviamente, a la mayor fuente de información que se posee: la descripción WSDL del servicio web. Orientada hacia los agentes de software que la interpretarán y utilizarán, esta descripción no siempre es simple de comprender. Es por este motivo que se construyen técnicas que agilicen y faciliten este proceso extrayendo la información adecuada para su comprensión [8,9,10].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo de investigación centra sus esfuerzos en la comprensión de servicios web cuyas descripciones son WSDLs. Esta comprensión se divide en tres grandes etapas, las cuales se describen sintéticamente en las subsecciones siguientes.

Extracción de información

La comprensión de los servicios web que se desea alcanzar hace necesaria la extracción de información de los mismos. En el caso particular de los servicios web, la información es bastante escasa. Se puede leer algo en los portales donde se publican los mismos como UDDI u alguna otra página de este tipo pero no siempre alcanza para comprenderlo. La mayor fuente de información de los mismos son sus descripciones WSDL. Por ser estas descripciones un dialecto XML, es fácil utilizar un parser DOM [11] que permita recorrerlos y generar árboles de sintaxis abstractas. Luego, desde estas representaciones en árboles se realizan recorridos que permiten extraer la información que se necesita. En esta investigación particular, se extraen los nombres de los identificadores de las etiquetas más importantes (*types, messages, portTypes, operation, bindings, services*), los comentarios y la documentación (etiqueta *documentation*) pero el árbol ya formado permite continuar extrayendo más información.

Análisis de la información y cálculo de métricas

Toda la información extraída es analizada para poder obtener una mensura de la comprensión que puede tener dicho servicio web a partir de su WSDL.

En el caso particular de los identificadores, se aplican algoritmos de división de palabras para aislar cada una de las palabras que lo componen, en caso de que existan, para luego analizar qué tanta semántica aportan.

De los comentarios y la documentación se realizan análisis semánticos de sus palabras más significativas, quitando las stopwords (palabras sin significado semántico) y enfocando el análisis en aquellas que aportan significado.

Sobre toda esta información obtenida se calculan métricas que ponderan este entendimiento parcial. Estas métricas son de dos tipos: i) de tamaño y ii) de semántica. Las primeras permiten conocer la dificultad del WSDL mediante el cálculo en las cantidades de cada una de sus etiquetas. Cantidad de mensajes, de tipos, de tipos complejos, de bindings, de portTypes son algunas de las métricas calculadas a partir de las cuales se puede obtener una idea rápida de la complejidad del servicio web y de cada una de sus partes. Las segundas métricas se encargan de valorar, por cada etiqueta del WSDL, la carga semántica que posee apoyándose en su nombre identificador, sus comentarios y su documentación. Para el análisis de esta carga semántica se utiliza NLTK [12,13] y una API de spelling, ambas apoyadas en el diccionario global WordNet [14] para analizar por cada palabra la carga semántica que posee y proveer sugerencias de palabras en caso que exista algún error o la palabra sea poco clara.

Grado de entendimiento del WSDL

Esta última etapa utiliza todos los valores de las métricas calculados en la etapa anterior para

cada etiqueta para calcular un grado de entendimiento global del WSDL. Basado en LSP [15], se construyó una estructura de agregación que relaciona todas las métricas semánticas otorgándole un peso a cada una de ellas y con distintos operadores propios de LSP con pesos en los arcos de la estructura, se van combinando los valores de las métricas para ir construyendo el grado de entendimiento total del servicio web. Por cuestiones de tamaño de este artículo, no se puede colocar la estructura de agregación completa pero a modo ilustrativo se muestra en la Figura 1 la estructura de agregación para el entendimiento de los mensajes.

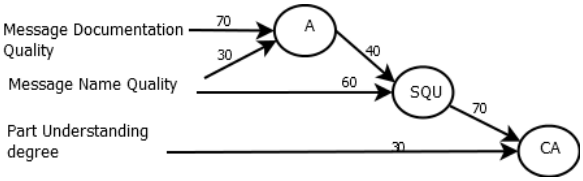


Figura 1. Estructura de agregación para medir el grado de entendimiento de los mensajes.

El lector interesado en profundizar más sobre estos conceptos puede leer [16].

Como conclusión a esta investigación se obtienen un conjunto de métricas que permiten calcular complejidades parciales por cada etiqueta principal del WSDL en tamaño y en comprensión semántica. Estas últimas sirven como base para la obtención del grado de entendimiento global del servicio web.

3. RESULTADOS OBTENIDOS/ESPERADOS

De los resultados obtenidos en esta investigación se mencionan los más destacados:

- Se definieron las estrategias para la extracción de información desde las especificaciones WSDL de los WS.
- Se realizó un trabajo manual de análisis de múltiples WSDL para poder determinar las claves para facilitar el entendimiento de los WSDL.
- Se definieron y establecieron todas las métricas necesarias, tanto de tamaño como semánticas.
- Se construyó la estructura de agregación LSP que permite vincular las métricas y calcular el grado de entendimiento global del WSDL.
- Se construyó una herramienta que realiza la extracción de información, cálculo de métricas, cálculo del grado de entendimiento del WS usando LSP (WSDLUD) y la visualización de la información. Esta herramienta permite además comparar múltiples servicios webs de manera automática mediante el cálculo de las métricas para cada uno de ellos.

Los objetivos planteados a corto y largo plazo son:

- Crear un módulo que permita recomendar automáticamente cambios en el WSDL para incrementar o decrementar el grado de entendimiento según sea necesario. Actualmente existe un trabajo de investigación avanzando con esta temática.
- Detectar y analizar automáticamente relaciones entre WSDLs mediante el estudio de sus espacios de nombres.
- Realizar un análisis más exhaustivo sobre los tipos complejos usados en cada WS.

4. 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 usando como base parte de esta investigación.

5. REFERENCIAS

[1] WSDL Specification for W3C https://www.w3.org/TR/wsdl.

[2] Erl, Thomas. Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services. Prentice Hall PTR. 2004. ISBN: 0131428985.

[3] Erl, Thomas. Service-Oriented Architecture: Concepts, Technology, and Design. Prentice Hall PTR. 2005. ISBN: 0131858580.

[4] Storey, M. "Theories, methods and tools in program comprehension: past, present and future," *13th International Workshop on Program Comprehension (IWPC'05)*, St. Louis, MO, USA, 2005, pp. 181-191. 2005.

[5] Von Mayrhauser, A. and Vans, A. M. "Program comprehension during software maintenance and evolution," In Computer, vol. 28, no. 8, pp. 44-55, Aug. 1995.

[6] N. Gold and K. Bennett. "Program comprehension for web services". In Program Comprehension, 2004. Proc. 12th IEEE International Workshop on. June 2004.

[7] 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.

[8] H. El Bouhissi, M. Malki, and D. Bouchiha. "A reverse engineering approach for the web service modeling ontology specifications". In Sensor Technologies and Applications2008. SENSORCOMM '08. Second International Conferenceon, pages 819–823, Aug 2008.

[9] C. Mao. "Towards a data complexity metric set for web service composition". In Computer and Information Tech-nology(CIT), 2011 IEEE 11th International Conference on, pages127–131, Aug 2011.

[10] 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.

[11] Parser DOM specification for W3C. https://www.w3.org/DOM.

[12] Natural Language Toolkit. . https://www.nltk.org/.

[13] Bird, Steven and Klein, Ewan and Loper, Edward. Natural Language Processing with Python. 2009. ISBN: 0596516495, 9780596516499. 1era Edición. O'Reilly Media, Inc.

[14] Wordnet, A Lexical database for English. https://wordnet.princeton.edu/

[15] Olsina, L., Rossi, G.: Measuring Web Application Quality with WebQEM. IEEE MultiMedia,2002 09(4), 20–29 (2002).

[16] 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.

MODELO DE ANÁLISIS DE FACTIBILIDAD DE IMPLANTACIÓN DE SOFTWARE

Marcela Dillon[1], Fernando López Gil[2], Mercedes Guash[2], Roxana Piergalini[2], Hugo Ramón[3], Balmer Agustin[4], Diaz Leonardo[4]

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, 236-4636945/44

mdillon@advantive.com.ar, {fernando.lopezgil, mercedes.guash, roxana.pergalini, hugo.ramon, agustin.balmer}@itt.unnoba.edu.ar

RESUMEN

A lo largo de las últimas décadas la ingeniería del software ha tenido una constante evolución en la definición del proceso de construcción de software. Sin embargo y pese a los grandes avances que ha producido, ésta se ha centrado principalmente en los problemas derivados del desarrollo, restándole atención a la implantación como parte integrante del proceso que coloca al producto desarrollado en un estado operativo.

Siguiendo la línea de investigación donde se desarrolló un modelo de implantación de software, se pretende ahora poder hacer un análisis de las condiciones del entorno donde se va a llevar a cabo la implantación y de ese modo poder detectar tempranamente potenciales dificultades en el proyecto a fin de tomar las medidas necesarias para evitarlas o al menos morigerarlas.

Palabras clave: Implantación de sistemas. Proceso software. Gestión de la implantación

CONTEXTO

Esta línea de investigación se realiza en el marco del proyecto "Tecnología y Aplicaciones de Sistemas de Software: Innovación en procesos", con lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) de la

---

1 Docente Investigador externo al ITT
2 Docente Investigador - ITT
3 Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC
4 Becario PROMINF - ITT
5 ITT - Centro Asociado CIC

Universidad Nacional del Noroeste de la Provincia de Buenos Aires.

Desde el año 2010 se ha estado investigando y desarrollando un modelo para la implantación de software que fue base para una tesis de maestría y publicado en el año 2016.

Siguiendo esta línea se continúa trabajando en el desarrollo de un modelo que permita la evaluación de las circunstancias en que se va a implantar el software, determine riesgos potenciales y acciones que sirvan para lograr su puesta en marcha de modo eficiente.

Se trabaja integrando docentes de grado y posgrado de la UNNOBA, alumnos, becarios e investigadores externos.

1. INTRODUCCIÓN

El proceso de Implantación es un tema relevante en lo que se refiere al desarrollo de software y de tecnologías de la información. A pesar de ello, la ingeniería de software continúa centrándose en abordar los problemas del desarrollo desde la mejora de procesos pero sin abordar de manera sistemática la implantación como un conjunto de temas específicos a ser tratados.

Uno de los problemas detectados en gran parte de los proyectos informáticos, en general está dado por las dificultades en la implantación de los mismos en los diferentes entornos sociales y tecnológicos, siendo esta etapa, un atributo fundamental para el éxito de dichos proyectos.

Si bien existen un sinnúmero de proyectos de software y estos a su vez tienen un sinnúmero de características que le son propias, se considera que se pueden determinar un cierto conjunto de características que son comunes para todos los proyectos de implantación.

Se busca encontrar por un lado, cuáles son los elementos que componen y afectan el proceso de implantación de software, no desde el punto de vista de un modelo que determine las actividades y procesos sino desde los componentes y las relaciones entre estos componentes que pueden afectar al normal desenvolvimiento de los procesos y actividades que son necesarias para la implantación.

Para hacerlo se propone trabajar en el grado de madurez de la organización, la complejidad del software a implementar –desde el punto de vista de la implantación-, el ambiente en que se encuentra la organización, las características del proyecto -grado de tangencialidad (término que pretende definir cuan tangente es al proceso de la organización o cuan amalgamado está, resultando esto en una menor o mayor dificultad al momento de la implantación) y otras características.

Por otra parte consideraremos a los actores del proceso, usuarios e implementadores.

Analizar en base a los factores planteados cuáles son los puntos principales que se deben estudiar para poder determinar las fortalezas, oportunidades, debilidades y amenazas que hay en cada etapa del proceso de implantación para cada proyecto.

En el proceso de implantación de software se deben considerar diferentes factores que pueden influir en el éxito –o fracaso- del mismo. Toda implantación se ve afectada entre otros por los siguientes factores:

- Madurez de la organización sobre la que se ha de implantar el software.

- Complejidad del software desde el punto de vista de la implantación, no desde su estructura algorítmica.
- Impacto del software en la organización.
- Ambiente en el que se encuentra la organización.
- Proyecto
- Usuarios
- Implementadores

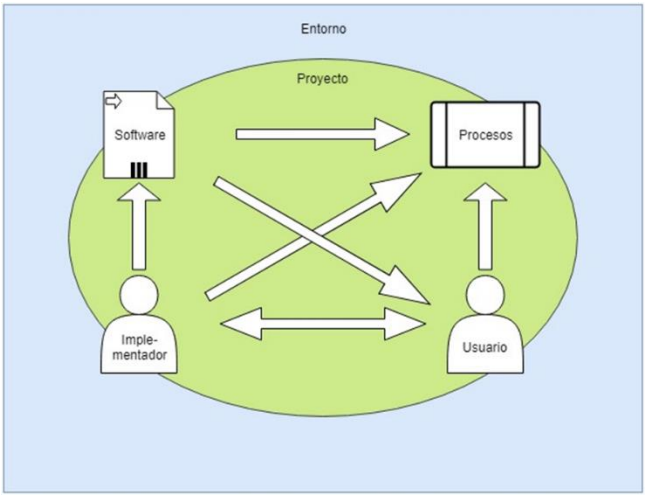


Figura 1 – Proceso de implantación

A su vez todos estos aspectos podrían ser analizados por separado, donde encontraremos diferentes factores que pueden darnos nuevas pistas en los puntos donde debemos prestar atención para lograr el éxito de nuestro proyecto.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La línea de investigación que se expone en el presente artículo, propone definir un Modelo de análisis de factibilidad para la implantación de software, basado en un conjunto de actores, artefactos, contexto y relaciones, que se constituya como una guía de actividades aplicables al proceso de puesta en marcha.

Para hacerlo se divide en dos áreas de trabajo: por un lado determinar una serie de criterios de evaluación de las circunstancias del proyecto que permitan detectar potenciales problemas en la implantación.

Por otro, poder determinar qué acciones o buenas prácticas son convenientes para aplicar en dichos proyectos de acuerdo a la evaluación realizada.

La línea que se plantea, está orientada a ser una línea de transferencia de tecnología a la industria, basada en el desarrollo de un modelo básico para abordar la Implantación de los sistemas, y es el resultado de una Tesis de Doctorado que se está desarrollando.

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo del proyecto es desarrollar un modelo que evalúe la factibilidad de implantación del software dadas las condiciones del proyecto y su entorno.

Como objetivos específicos se pretende:

- Definir el conjunto de elementos del proceso de implantación que puedan ser fácilmente evaluables y que permita definir características que influyan en el proceso.
- Construir una matriz que permita determinar la factibilidad del proyecto según los elementos evaluados.
- Determinar las acciones adecuadas para la implantación según la evaluación de la factibilidad del proyecto.

- Construir un framework de soporte al modelo.

4. FORMACIÓN DE RECURSOS HUMANOS

Se pretende propiciar un marco formal con el establecimiento de esta línea de investigación, para atender a la formación de recursos humanos, en su rol de investigadores o partícipes activos en el equipo de investigación; fomentando así la culminación de sus estudios superiores, y promoviendo la redacción, exposición y defensa de Trabajos Finales de Grado y Postgrado.

También en el término de tres años, se espera contribuir al inicio y concreción de 1 (una) Tesis de Magíster y la finalización de 1 (una) Tesis Doctoral.

Con la incorporación de becarios y alumnos al proyecto, se busca desarrollar futuros investigadores en el área de ingeniería de software dentro de la universidad.

5. BIBLIOGRAFÍA

Church, J. W. "Organization Development: a Data-Driven Approach to Organizational Change". San Francisco: CA: Published by Jossey Bass. (2001)

CMMI. "Capability Maturity Model® Integration Version 1.3. CMMI-DEV for Systems Engineering, Software Engineering, Integrated Product and Process Development, and Supplier Sourcing. (CMMI-SE/SW/IPPD/SS,V1.3)". Carnegie Mellon University. Software Engineering Institute, USA, (2013)

Chrissis, Mary Beth; Konrad, Mike; Shrum, Sandy. "CMMI® Guía para la integración de procesos y la mejora de productos". (2009)

Mon, Alicia; López Gil, Fernando. "Modelo de Implantación de Software - ISM". Editorial: Grupo Editor TM. ISBN 978-987-4137-01-2 (2016)

Mon, Alicia; López Gil, Fernando. "Implantación de Software, un Modelo Básico". WICC 2014. (2014)

Modelo de Análisis para la derivación de requerimientos funcionales a partir de la implementación de Patrones en la construcción del Modelo Conceptual

Claudia Castro, Andrea Delgado, Marcelo Marciszack, Juan Carlos Moreno
Dpto. Ingeniería en Sist. de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
{ingclaudiacastro, andreafdelgado, marciszack, jmoreno33}@gmail.com

RESUMEN

El proyecto tiene como objetivo caracterizar un Modelo de Análisis que permita definir requerimientos en la actividad de Modelado Conceptual de un sistema de información, usando Patrones para el análisis de procesos de negocio.

En base a los Patrones de Negocio, se propondrán diferentes técnicas para optimizar la selección de patrones que serán de utilidad para el modelado de la vista interna del proceso de negocio bajo análisis, es decir, la especificación detallada de las actividades que comprenden el mismo. De igual modo, se establecerán un conjunto de buenas prácticas para la definición de requerimientos funcionales dentro de la fase de construcción de un modelo conceptual de un sistema de información, partiendo de la vista interna definida; en otras palabras, se determinará el proceso de derivación de los requerimientos funcionales del sistema de información.

CONTEXTO

Este proyecto, puede considerarse como una extensión al proyecto PID EIUTNCO00066604; Implementación de patrones en la validación de Modelos Conceptuales, consolidado dentro del Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información (CIDs). Se incorporaron en el Modelado Conceptual los patrones de negocio, precisando la especificación detallada del proceso de negocio, que serán utilizados para facilitar la definición inicial de un nuevo sistema de información.

1. INTRODUCCIÓN

Existen una gran variedad de trabajos que evidencian que los errores que se cometen en la etapa de especificación de requerimientos, para la obtención de un esquema conceptual, tienen un costo relativo en relación a su reparación y crecerá en forma exponencial a medida que se avanza en etapas. Hay trabajos que describen la importancia de los requerimientos y el gerenciamiento de los mismos; basándose en la definición de [Ackoff 1974] "Fallamos más a menudo porque resolvemos el problema incorrecto, que porque obtenemos una solución deficiente al problema correcto". "Nuestro fracaso en dominar la complejidad nos lleva a proyectos rechazados, que exceden los presupuestos y que son deficientes con respecto a los requerimientos fijados". La preocupación por definir los requisitos de manera adecuada es extensamente tratada en [Sommerville 2011], donde el eje central es la definición de buenas prácticas en la creación de los mismos, en donde se 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 homogéneo y bien entendido que el proceso de desarrollo de software en su conjunto".

El objetivo final, es crear el sistema de software (s) que mejor soporte y se ajuste al negocio; entonces, el modelo de negocio, es muy importante tanto para especificar requisitos como para diseñar software.

Para modelar el negocio se cuenta con diferentes herramientas que facilitan, guían y estandarizan la labor. En tal sentido, es necesario conocer conceptos tales como: BPM, Patrones de modelado de Negocio, Patrones de modelado de Procesos de Negocio, UML, entre otros.

Los responsables de sistemas de información entienden a la "Arquitectura empresarial" como la descripción de los principales elementos y relaciones que hacen posible a una organización, entendiendo como parte esencial de estos elementos a las aplicaciones informáticas y a los procesos de negocio. La alineación de estas partes son las que llevarán al cumplimiento de los objetivos del negocio, y los procesos de negocio los que podrán mejorar a través de la tecnología. A partir del análisis de los procesos de negocio, de sus partes y sus pasos, es que se puede pensar en su interacción con el sistema de información. Los casos de uso constituyen una herramienta que permitirá capturar los requerimientos funcionales a partir de los pasos de un proceso de negocio, y los actores a partir de los roles participantes en el proceso. Los requerimientos que afectarán al sistema de información, también pueden derivarse desde el análisis de cada uno de los procesos de negocio, en definitiva son las características deseadas del proceso aquellas que determinan aspectos funcionales. Al análisis de procesos de negocio, se debe sumar el análisis de la información. ¿Y qué información es relevante para nuestro caso? Aquella que modela aspectos esenciales y conceptos del negocio, entendiendo a los datos como el recurso del sistema de información, que se procesa para brindar información sobre y para los procesos de negocio.

Se utilizan patrones debido a que son reusables, simples, probados, genéricos y proveen calidad al desarrollo; esto permite reutilizar no solo modelos, sino también componentes y piezas completas de software y la arquitectura que les da soporte y, a su vez, establecen un lenguaje común para la solución de problemas específicos, volviéndose fácil de entender y manejar. Se buscan patrones que sean aplicables a un problema específico, dentro de un contexto determinado y den soluciones claras sobre el problema planteado.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Para la consecución de los objetivos del presente proyecto se seguirá el método científico, en donde se focalizará la validez de la propuesta de trabajo en base al grado de aceptación y utilización que se le añadirá a la técnica propuesta en diferentes proyectos de transferencia que se realizan dentro del CIDS. Para alcanzar los objetivos particulares se aplicarán el siguiente conjunto de actividades:

1) Especificación del dominio de estudio y alcances de la investigación.

Evaluar patrones de negocio a partir del análisis de sus características, efectuando una búsqueda bibliográfica de documentación pertinente y artículos relacionados al tema.

2) Conceptualización del Patrón.

Completar el estudio de las distintas estructuras y clases de patrones de negocio, analizando las características básicas de los patrones identificados. Validar la aplicación de los patrones utilizando los mismos en la construcción de Modelos Conceptuales para dominios planteados en diferentes casos de estudio, tanto académicos como reales. Analizar herramientas y métodos aplicables al análisis y la especificación detallada de los procesos de negocio, estableciendo sus características. Investigar, caracterizar y especificar métodos y metodologías, así como herramientas, técnicas, patrones, lenguajes y notaciones utilizados para el análisis y especificación de Procesos de Negocio (tales como BPMN, UML, Diagrama de Flujo, Cursogramas, IDEF, y otras) que faciliten la construcción del Modelo Conceptual, con sus ventajas y desventajas; en base al juicio empírico de profesionales y docentes del área.

3) Análisis de los patrones en entornos de negocio y sistemas de información.

- Determinar las herramientas y métodos de análisis de procesos de negocio, que pueden

colaborar con la identificación de requerimientos del sistema de información de soporte. Aplicar los métodos, herramientas y patrones especificados en el objetivo 2, con el fin de divisar los aspectos esenciales para el proceso de derivación de requerimientos. Determinar aquellos que se adecúan a la derivación de requerimientos a partir de diferentes Modelos Conceptuales.

4) Proposición de un nuevo marco de derivación de requerimientos a partir del modelo de negocios para el modelado conceptual de sistemas.

Plantear un conjunto de estrategias de derivación de requerimientos a partir del modelado de los procesos de negocio y de la arquitectura del negocio. Determinar implicancias de derivar requerimientos de un sistema de información. Analizar alternativas a partir de los métodos, herramientas y patrones analizados en el objetivo 3, a fin de establecer un conjunto de estrategias, buenas prácticas y recomendaciones que permitan definir los requerimientos del sistema de información. Obtener los requerimientos que dan soporte a los procesos de negocio basados en un Modelo Conceptual.

5) Validación y verificación del marco metodológico propuesto.

Confirmar el marco metodológico propuesto, empleando casos de estudio y casos reales.

6) Diseño y construcción de herramienta.

Desarrollar una herramienta gráfica que permita:

- Capturar el modelo de proceso de negocio generado a partir de la aplicación de patrones de proceso de negocio.
- Representar las actividades del proceso de negocio, bajo análisis, mediante herramientas gráficas (por ejemplo: BPMN, IDEF).
- Determinar actividades manuales y de software según criterio de expertos.
- Clasificar y segregar las actividades identificadas.
- Obtener el listado de requerimientos de manera nominal.
- Generar y agrupar los requerimientos según criterio de expertos.

7) Difusión de conocimientos adquiridos.

Utilizar y divulgar el conocimiento adquirido/desarrollado, es decir, preparar artículos y seminarios, para validar con la comunidad científica el avance de la propuesta, y generar divulgaciones sobre el tema, en forma interna y externa. Desarrollar y ejecutar planes de capacitación y asistencia a usuarios finales para transferencias efectivas tanto académica como organizacionalmente.

3. RESULTADOS ESPERADOS

Caracterizar un Modelo de Análisis que permita definir requerimientos funcionales en la actividad de Modelado Conceptual de un sistema de información empleando Patrones durante la etapa de análisis de procesos de negocio. Objetivos específicos:

1) Evaluar patrones de negocio existentes a partir del análisis de sus características.

2) Analizar herramientas y métodos aplicables al análisis y la especificación detallada de los procesos de negocio, estableciendo sus características.

3) Determinar las herramientas y métodos de análisis de procesos de negocio que pueden colaborar con la identificación de requerimientos del sistema de información de soporte.

4) Proponer un conjunto de estrategias de derivación de requerimientos a partir del modelado de los procesos de negocio y de la arquitectura del negocio.

5) Validar la derivación de los requerimientos, obtenidos a partir del modelo de negocio.

6) Desarrollar una herramienta gráfica que dé soporte a la derivación de requerimientos, basado en las estrategias propuestas.

7) Utilizar y divulgar el conocimiento adquirido/desarrollado.

4. FORMACIÓN DE RECURSOS HUMANOS

La ejecución de este proyecto prevé la formación de los integrantes que forman parte del mismo, a saber: Tesis doctoral dentro del Doctorado en Ingeniería: mención Sistemas de Información del tesista Mgter. Ing. Juan

Carlos Moreno cuyo tema es “Metodología de la evaluación temprana de la usabilidad empleando patrones en la construcción del modelo conceptual de aplicaciones web”. Tesis doctoral dentro del Doctorado en Ingeniería: mención Sistemas de Información cuyo tema es “Definición de patrones a partir de buenas prácticas para el desarrollo de sistemas de Gobierno Electrónico”, en ejecución por el Ing. Oscar Carlos Medina. Además, se incorporaron al equipo de trabajo, docentes-investigadores de la carrera Ingeniería en Sistemas de Información como investigadores de apoyo; se incorporará a dos becarios alumnos de investigación y un becario Graduado BINID, que serán de ayuda en la recolección y manipulación, y colaborarán en el desarrollo de este marco metodológico.

5. BIBLIOGRAFÍA

[Ackoff 1974] Ackoff, R. L., “Redesigning the future”, New York, (1974).

[Antonelli 2003] Antonelli, L., Oliveros, A., Rossi, G., “Traceability en la Elicitación y Especificación de Requerimientos”, Tesis de Maestría Universidad Nacional de la Plata, (2003).

[Barrett 2006] Barrett, R., Patcas, L.M., Pahl, C. & Murphy, J.. “Model driven distribution pattern design for dynamic web service compositions”. in Proceedings of the 6th international conference on Web engineering, págs. 129-136, ACM, (2006).

[Beck 1996] Beck, D.; Coplien, J.; Crocker, R.; Dominick, L.; Meszaros, G.; Paulisch, F. Industrial experience with design patterns. ICSE?18 (International Conference on Software Engineering), Technical University of Berlin, Germany. Pág.: 103?113.

[Boehm 2001] Boehm, Barry, Basili, Victor R.: “Software defect reduction top 10 list”, IEEE Computer, Vol. 34, N° 1, (2001).

[Booch 1996] Booch, G., Jacobson, I., & Rumbaugh, J. The unified modeling language. Unix Review, 14(13), 5. (1996).

[Braude 2010] Braude, E.J., & Bernstein, M., “Software Engineering: Modern Approaches”, 2nd ed. Wiley, (2010).

[Carrasco 2010] Gestión de Procesos (La Participación es la clave). Juan Bravo Carrasco, Editorial Evolución S.A. – 2010. [Eriksson, Penker 2000] Busines Modeling with UML Business Patterns at Work; Hans-Erik Eriksson. Magnus Penker 2000.

[Foerster 2005] Foerster, Alexander; Engels, Gregor; Scharokowsky, Tim. Activity diagram patterns for modeling quality constraints in business processes. En Model Driven Engineering Languages and Systems. Springer Berlin Heidelberg, 2005. p. 2?16. [Fowler 1997] Fowler, M., Analysis Patterns: Reusable Object Models, Addison Wesley. 1997.

[Gamma 1995] Gamma, E., Helm, R., Johnson, R., & Vlissides, J. Design patterns: elements of reusable object?oriented software. Addison?Wesley, 1995.

[Imamura 2005] Imamura, T., Tatsubori, M., Nakamura, Y., “Web Services Security Configuration in a Service-Oriented Architecture”, IBM Research, Tokyo Research Laboratory, 1623-14, Shimotsuruma, Yamato, Kanagawa 242-8502, Japan, and Christopher Giblin, IBM Research, Zurich Research

[Insfrán 2002] Insfrán, E., Díaz, I. y Burbano, M., “Modelado de Requisitos para la Obtención de esquemas conceptuales”, (2002).

[Judson 2003] Judson, S.R.., Carver, D.L.., France, R.B., “Pattern-Based Model Transformation”, Computer Science Department, Louisiana State University, Baton Rouge, Louisiana USA. OOPSLA’03, October 26–30, 2003, Anaheim, California, USA. ACM 1- 58113-751-6/03/0010, (2003).

[Larman 2003] UML y Patrones 2da. Edicion, Cairg Larman, 2003.

[Maassen 2007] Maassen, Olaf; Stelting, Stephen. Creational Patterns: Creating Objects in anOO System, 2007.

[Marciszack 2015] Marciszack, M., Perez Cota, M., and Groppo, M., "Metodología y Herramienta de soporte para validar Modelos

Conceptuales a través de Máquinas Abstractas", Revista de Ciencia y Tecnología, U. de Palermo, Nro 15, págs. 165-180. ISSN: 1850-0870 ISSN (en línea): 2344-9217. 2015

[Marciszack 2016] Marciszack, M., Castro, C., Sánchez, C., Delgado, A., Garnero, A. B., Horenstein, N., Fernández, E. "Una experiencia en la aplicación de Patrones de Negocio". IV CONGRESO NACIONAL DE INGENIERÍA EN INFORMÁTICA/SISTEMAS DE INFORMACIÓN, CoNaIISI 2016, Publicación on line - ISSN 2347-0372. 2016.

[Marciszack 2017] Marciszack, Marcelo, Castro, Claudia, Delgado, Andrea, Sánchez, Claudia, Horenstein, Nicolás, Garnero, Ana Belén. "Aplicación y Uso de Patrones para el Modelado de Procesos de Negocios en el ámbito educativo y sector gubernamental". En Tecnología, Innovación y Creatividad: III JATIC 2017 (pp. 132-136) compilado por Malbernat, L.R., Finochietto, J.R, Cormons, M.A., Varela, A.E. Universidad CAECE, Mar del Plata, Argentina. 2017.

[Medina 2015] Medina, O., Marciszack, M., Groppo, M. "Herramienta para administración y validación de requerimientos de sistemas". 4th CIMPS 2015, "Tendencias en la Ingeniería del Software. Impacto en las Tecnologías de Información y Comunicación", Editorial: CIMAT A.C., Mazatlán, Sinaloa, México, ISBN: 978-607-96212-6-1. 2015.

[Medina 2016] Medina, O., Marciszack, M., Groppo, M., "Trazabilidad y validación de requerimientos funcionales de sistemas informáticos mediante la transformación de modelos conceptuales". Revista ReCIBE, Año 5 No. 1, Universidad de Guadalajara. Guadalajara, Jalisco, México, 2016.

[Penker 2000] Eriksson, Hans-Erik; Penker, Magnus. Business modeling with UML. New York, p. 1-12. (2000).

[Pons 2010] Pons, C., Giandini, R. S., Pérez, G. "Desarrollo de software dirigido por modelos. Conceptos teóricos y su aplicación práctica". Editorial de la Universidad Nacional de La Plata (EDULP) / McGraw-Hill Educación, Facultad de Informática, La Plata, Buenos Aires. (2010). ISBN: 978-950-34-0630-4.

[Pressman 2005] Pressman R., Jesús Elmer Murrieta, Ingeniería del software: un enfoque práctico, 6°Ed., McGraw?Hil, Interamericana, España, 2005.

[Pressman 2010] Pressman, R.S., "Ingeniería del Software. Un enfoque práctico", 7ma. Edición, Editorial Mc Graw Hill Educación - México (2010).

[Rahmani 2006] Rahmani, A.T., Rafe, V., Sedighian, S., Abbaspour, A., "An MDA-Based Modeling and Design of Service Oriented Architecture.", Iran University of Science and Technology, Computer Engineering Department, Tehran, Iran. V.N. Alexandrov et al. (Eds.): ICCS 2006, Part III, LNCS 3993, pp. 578 – 585, 2006. © Springer-Verlag, Berlin Heidelberg, Germany, (2006).

[Räihä 2009] Räihä, O., Mäkinen, E., Poranen, T., "Using Simulated Annealing for Producing Software Architectures.", University of Tampere, Department of Computer Sciences, FIN-33014 University of Tampere, Finland. GECCO'09, July 8–12, 2009, Montréal Québec, Canada ACM 978-1-60558-505-5/09/07, (2009).

[Sánchez 2013] Sánchez, C. "Modelado de Procesos de Negocio con Aplicación de Patrones". Trabajo final de la Especialidad en Ingeniería en Sistemas de Información, Biblioteca de PostGrado de la Universidad Tecnológica Nacional, Facultad Regional Córdoba.

[Sommerville 2011] Sommerville, Ian. "Ingeniería de Software", 9° Ed., Pearson Educación, México, 2011, ISBN: 978?607?32?0603?7.

[Sesé 2007] Sesé Muniátegui, F., "Propuesta de un método de validación de esquemas conceptuales y análisis comparativo de la noción de información en los métodos de desarrollo de Sistemas de información", Tesis Doctoral Universitat Ramon Llull, España (2007).

[Verdanat 1996] Verdanat, FB. "Enterprise modeling and integration: principles aapplication." (1996).

Modelos y Métodos de Calidad: Fortalecimiento de la Seguridad en los Sistemas de Software

Carlos Salgado, Mario Peralta, Mario Berón
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, mberon}@unsl.edu.ar

RESUMEN

Los avances tecnológicos aportan un nuevo nivel de autonomía a los vehículos, los robots en los almacenes, las cámaras de seguridad y una amplia gama de servicios de Internet. Por lo tanto, los autos, datos personales, métodos de pago electrónico, servicios en la nube y dispositivos personales de empleados serán algunos de los blancos de los *ciberdelincuentes*. La seguridad se ha convertido en un punto crítico para la vida diaria de las personas. Dicho avance, nos lleva a combinar y extender las características deseables o esperables de los recursos humanos para garantizar la seguridad de datos e información de las organizaciones. El incremento en la seguridad, con el uso de modelos de calidad adecuados, con posibilidad de implementar un tablero de control, asegura un adecuado resguardo de toda la información sobre los proyectos en marcha, que es esencial para poder gestionar de manera eficiente.
En este sentido, la presente propuesta consiste en la definición de un modelo que se basa en las mejores características de seguridad de la Norma ISO 25000, tales como Confidencialidad, Integridad, Responsabilidad, Autenticidad y las consideraciones de la ISO 9001 con respecto a las condiciones o características de los recursos humanos de las organizaciones.

Palabras Claves: Seguridad, Ataques al Personal, Métodos de Evaluación, Modelos de Calidad,

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 previsiones para 2020 indican que habrá, al menos, 200.000 millones de dispositivos conectados a la red. Y no serán solo computadoras o teléfonos, sino que habrá electrodomésticos, autos, relojes, ropa, entre otros, y serán los blancos para los piratas informáticos. Estas consideraciones ponen de manifiesto la necesidad de modelos, métodos, técnicas y herramientas para poder combatir, o al menos prevenir, dichos ataque. Hoy en día, el acceso a los datos e información está dado desde distintos aparatos o dispositivos electrónicos.
Internet es el medio que permite tener alcance a mayores distancias y diversidad de actores.
Las aplicaciones web son utilizadas en una amplia variedad de áreas, entre ellas, redes sociales, compras, actividad bancaria, sistemas de control, almacenamiento en la nube y demás. Inclusive desde el año 2011, el 75 % de

la distribución de aplicaciones por categorías correspondía a aplicaciones web. Este abrupto crecimiento se debe, entre otras cosas, a la versatilidad que supone un navegador web como cliente ligero, la independencia con el sistema operativo y la posibilidad de mantener o actualizar la aplicación sin la necesidad de interferencia de los usuarios potenciales [1].

Para intentar satisfacer la incesante demanda de sistemas de este tipo, los usuarios ocasionalmente se ven rodeados de software que no fue desarrollado siguiendo un modelo adecuado, que no ha sido codificado siguiendo parámetros de calidad y/o posee un diseño defectuoso, entre otros tantos factores que vuelven a los sistemas poco robustos y con grandes falencias en seguridad [2].

En un contexto de conectividad permanente entre los usuarios, el concepto de "seguridad" se vuelve un factor fundamental [3]. Este motivo, entre otros, es el principal desencadenante de una disciplina que se encuentra en auge en distintos ámbitos de la actualidad, la misma es conocida como: Seguridad Informática.

La noción de seguridad en los Sistemas de Información, se define como un grupo de componentes con características especiales, de los cuales se pueden destacar [4]:

Seguridad: incluye características en concomitancia con la protección del sistema, sus aplicaciones y los recursos compartidos. Incluye la prevención de la adquisición y modificación no autorizada de información, es decir, se intenta cubrir tanto la seguridad del sistema como la de los datos del usuario.

Fiabilidad, disponibilidad y recuperación: indican, respectivamente, la exactitud de la seguridad del sistema y otras funciones del mismo, el soporte a través del tiempo de un sistema seguro y la capacidad de recuperación del mismo luego de haber sufrido un ataque o algún tipo de error/accidente.

Auditabilidad: implica el seguimiento de la continua existencia de seguridad y fiabilidad del sistema, incluyendo la detección de anomalías o un amenazante comportamiento.

Un camino viable para conseguir, como producto final, un sistema que provea mecanismos de seguridad adecuados, es seguir un enfoque correctivo [5]. Esto quiere decir, llevar a cabo la evaluación de la seguridad del software a posteriori de su desarrollo. En este sentido, es preciso hacer referencia a un conjunto de herramientas denominadas: "Herramientas de Análisis de Seguridad de Software".

La seguridad es un atributo de calidad del software y, como parte de un todo, al llevar a cabo un análisis, tradicionalmente se distinguen dos tipos: el *Análisis Dinámico*, donde se diseñan casos de prueba sobre la información base de la especificación, y el *Análisis Estático*, donde las técnicas de evaluación o prueba se diseñan en base a la información derivada del código fuente [6, 7].

Más allá del camino que se tome para realizar sus procedimientos, el fin de las herramientas de análisis de seguridad es encontrar aquellos puntos donde el sistema informático sufre problemas relacionados o, inclusive, identificar aquellas vulnerabilidades que podrían ser potencialmente conflictivas. De acuerdo con el Instituto SANS [8], en el ámbito de aplicaciones web, las herramientas específicas de análisis de seguridad se pueden clasificar, teniendo en cuenta la forma que utilicen para llevar a cabo sus procesos, en al menos cuatro categorías: i) Bloqueo de ataques: basados en la red de datos. ii) Bloqueo de ataques: basados en el servidor de host. iii) Eliminación de vulnerabilidades de seguridad. iv) Soporte seguro para usuarios autorizados.

De esta forma, se puede concluir que existe una gran variedad de posibilidades a la hora de seleccionar una herramienta que ejecute un proceso de análisis de seguridad para aplicaciones Web. Dentro de lo que significa la detección y eliminación de vulnerabilidades de seguridad, con el fin de acotar el alcance de la investigación y, por lo tanto, conseguir un enfoque más preciso, este artículo se centra en la evaluación de los recursos humanos que utilizan las aplicaciones web en las

organizaciones/instituciones.

Desde otro punto de vista, el ataque a través de los empleados es un punto de vulnerabilidad cada vez expuesto debido a las distintas posibilidades de acceso que aportan las nuevas tecnologías. Las organizaciones continuarán mejorando sus posturas de seguridad, implementando las últimas tecnologías de seguridad, trabajando para contratar a personas con talento y experiencia, creando políticas efectivas y permaneciendo vigilantes. Sin embargo, los atacantes probablemente cambien su enfoque y ataquen cada vez más a las empresas a través de sus empleados, dirigiéndose entre otras cosas, a los relativamente inseguros sistemas del hogar de los empleados para acceder a las redes corporativas.

Desde este punto de vista, la norma ISO 9001-2015 [9] trata sobre el recurso humano de una organización. Para que ésta pueda satisfacer a sus clientes con productos de calidad, debe tener un personal de calidad. El recurso humano se considera de calidad cuando:

- es competente en base a cuatro aspectos: educación, formación, habilidades y experiencia;
- está consciente de la importancia de sus actividades en relación con la calidad, y
- está satisfecho.

De acuerdo a la Norma ISO 9001-2015, la organización debe asegurar que su personal es *consciente* de la relevancia e importancia de sus actividades y cómo ellas contribuyen a la consecución de los objetivos de calidad.

Basados en los principios de participación del personal y de la orientación a procesos, tener un personal consciente es una labor de formación y de creación de una cultura organizacional. Que debe garantizar la Alta Dirección de la organización al asegurar que las responsabilidades y autoridades son definidas y comunicadas dentro de la organización.

Por otro lado, el personal debe ser *competente*. Entendiendo competencia como un "conjunto de comportamientos observables que están causalmente relacionados con un desempeño bueno o excelente en un trabajo concreto y en una organización concreta" [10].

La norma ISO 9001 considera que una persona es competente cuando cumple los requisitos de educación, formación, habilidades y experiencia que la organización determina para cada puesto de trabajo. Donde:

- Educación: Estudios mínimos requeridos para un determinado puesto.
- Formación específica: Son todos aquellos conocimientos adicionales que necesarios para desempeñar las actividades de un puesto, por ejemplo, especialización en determinadas herramientas informáticas.
- Habilidades especiales de tipo práctico. Por ejemplo, la habilidad que debe tener un vendedor. Estas habilidades, específicas para cada puesto, facilitan el desempeño del mismo.
- Experiencia mínima, que el trabajador debe tener en el puesto o en puestos similares y que incluye un período de prácticas mínimo en la empresa.

Estas competencias son de tipo umbral, es decir, las mínimas que debe tener una persona para realizar su trabajo con éxito, pero no van a diferenciar a los trabajadores con desempeño excelente de los que tienen un desempeño normal.

En cuanto a la *satisfacción,* un personal satisfecho es aquel personal motivado, que siente que tiene cubiertas sus necesidades personales, que se siente a gusto con el trabajo que desempeña y estima que tiene estabilidad en él. Tiene la tranquilidad suficiente como para dedicar todos sus esfuerzos físicos, psíquicos e intelectuales a desarrollar su trabajo sin que tensiones o preocupaciones ajenas a las actividades propias del puesto lo distraigan.

Un elemento que contribuye a una buena productividad y a la satisfacción del empleado es el mantenimiento del área de trabajo en condiciones cómodas y saludables. En la Norma ISO 9001-2015 se define que el ambiente de trabajo está relacionado con aquellas condiciones en las cuales se realiza el trabajo incluyendo los factores físicos, ambientales o de cualquier otro tipo. En este punto, la norma establece que la organización debe determinar y gestionar el ambiente de

trabajo necesario para lograr la conformidad con los requisitos del producto. En este sentido es importante el estudio de los aspectos ergonómicos y de seguridad e higiene en el área de trabajo.

De ahí que las organizaciones deben intentar controlarlos y crear condiciones lo más cómodas y seguras posibles para que el trabajador pueda ser eficiente sin que se vea afectado por estos elementos. Realizar un trabajo en condiciones incómodas o inseguras crea en el trabajador un sentimiento de descontento pues le hace pensar que la empresa no se preocupa por cómo realiza su trabajo.

Desde este punto de vista, al contar con un modelo y/o método se debe procurar prevenir o anticiparse para lograr mantener los procesos organizacionales funcionando de forma segura y sistemática para garantizar la satisfacción y trabajo de calidad del personal de las organizaciones. Ese es el actual desafío que presenta el nuevo paradigma de negocio y las bondades de las TIC.

En base a ellos se analizan los beneficios de definir un modelo y/o método de calidad que permita establecer las pautas y pasos a seguir a la hora de analizar la existencia de vulnerabilidades de seguridad en las aplicaciones web, las cuales son el nexo entre los empleados y su trabajo. Tal vez considerar los nuevos paradigmas de trabajo a distancia y lo que conlleva al acceso a los recursos de la empresa/organización de manera remota.

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.
- Hacer una investigación profunda sobre las vulnerabilidades de seguridad que afectan a los sistemas de Información.
- Investigar el estado del arte de los patrones/estilos arquitectónicos de sistemas de información, y de los protocolos utilizados en su desarrollo.
- Integrar las investigaciones realizadas en ambos proyectos con el fin de crear un conjunto de herramientas que faciliten a los distintos actores en ciclo de desarrollo de software poder interactuar en pos de la seguridad de los sistemas, recursos de la organización.
- Análisis de los distintos protocolos de seguridad. Pensando en la definición de reglas, normas, protocolos, guías que permitan desarrollos 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

Según lo establece la Norma 25010 [11], la seguridad se define como la "Capacidad de protección de la información y los datos de manera que personas o sistemas no autorizados no puedan leerlos o modificarlos", y, considerando las subcaracterísticas esperables o deseables para un producto software en cuanto a seguridad, el modelo que se propone en esta línea de investigación consiste de las siguientes características:

- *Confidencialidad*. Capacidad de protección contra el acceso de datos e información no autorizados, ya sea accidental o deliberadamente.
- *Integridad*. Capacidad del sistema o componente para prevenir accesos o modificaciones no autorizados a datos o programas de ordenador.
- *No repudio*. Capacidad de demostrar las acciones o eventos que han tenido lugar, de manera que dichas acciones o eventos no

puedan ser repudiados posteriormente.

- *Responsabilidad*. Capacidad de rastrear de forma inequívoca las acciones de una entidad.
- *Autenticidad*. Capacidad de demostrar la identidad de un sujeto o un recurso.

Además de estas características, se considera de suma importancia la subcaracterística: ***ataques a través de los empleados***, la cual tiene que ver con lo que se refiere al personal dentro de una organización. Las organizaciones continuarán mejorando sus posturas de seguridad, implementando las últimas tecnologías de seguridad, trabajando para contratar a personas con talento y experiencia, creando políticas efectivas y permaneciendo vigilantes. Sin embargo, los atacantes probablemente cambien su enfoque y ataquen cada vez más a las empresas a través de sus empleados, dirigiéndose entre otras cosas, a los relativamente inseguros sistemas del hogar de los empleados para acceder a las redes corporativas. Por lo tanto, otra arista más a considerar es la seguridad aplicada fuera de la organización.

Es por ello que se está trabajando en la definición del modelo mencionado. El objetivo es proveer una herramienta que permita prevenir en gran medida los distintos ciberataques a los que los usuarios de las nuevas tecnologías están expuestos en la actualidad.

4. FORMACION 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. Además, en el marco de esta línea de investigación, se está desarrollando una tesis para la Maestría en Calidad de Software dictada por la Universidad de San Luis bajo Resolución ME N° 1589/13. En el ámbito de dicho proyecto se propone la investigación sobre modelos y métodos de calidad que fortalezcan el desarrollo de software.

5. BIBLIOGRAFÍA

[1] R. CHOPRA. Web Engineering. PHI Learning Pvt. Ltd., 2016.

[2] P. E. Black. Software assurance metrics and tool evaluation. In Software Engineering Research and Practice, pages 829–835, 2005.

[3] C. Vulnerabilities. Exposures,“the standard for information security vulnerability names”. Common Vulnerabilities and Exposures: The Standard for Information Security Vulnerability Names. url: http://cve. mitre. org, 2007.

[4] J. Song, G. Hu, and Q. Xu. Operating system security and host vulnerability evaluation. In Management and Service Science, 2009. MASS’09. International Conference on, pages 1–4. IEEE, 2009.

[5] P. G. Neumann. Computer system security evaluation. In National Computer Conference, 1978.

[6] M. Ishrat, M. Saxena, and D. M. Alamgir. Comparison of static and dynamic analysis for runtime monitoring. International Journal of Computer Science & Communication Networks, 2:615–617, 2011.

[7] C. Artho and A. Biere. Combined static and dynamic analysis. In Proc. AIOOL ’05, ENTCS, pages 98–115. Elsevier Science, 2005.

[8] I. SANS. Seguridad 101: Los tipos de malware. https://www.sans.org/critical-security-controls/, 2017.

[9] ISO ORG. (Septiembre de 2015). *ISO 9001:2015*. Recuperado el octubre de 2016, de http://www.iso.org/iso/home/standards/management-standards/iso_9000.htm

[10] Pereda, S., Berrocal, F., Sanz, P. (2003), Los perfiles de exigencias en la ocupación del profesional de recursos humanos, Psicología desde el Caribe, Universidad del Norte, Nº 12

[11] ISO/IEC 25010:2011. Systems and software engineering -- Systems and software Quality Requirements and Evaluation (SQuaRE) -- System and software quality models."

FOPTIMIZACIÓN MULTIOBJETIVO BASADA EN PREFERENCIAS PARA PROBLEMAS DE LA INGENIERÍA DE SOFTWARE

Carlos Casanova, Martín Arrúa, Luciano Bracco, Fernando Pereyra, Giovanni Daián Rottoli, Esteban Schab, Adrián Tournour, 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
{casanovac, arruam, braccol, pereyraf, rottolig, schabe, tournoura, debattistaa}@frcu.utn.edu.ar

RESUMEN

La Ingeniería de Software Basada en Búsqueda (ISBB) se sustenta en la aplicación de técnicas de optimización heurística para resolver problemas de la Ingeniería de Software (IS). Las tarea de la IS se enmarcan en un problema de búsqueda, con lo cual es posible utilizar multitud de algoritmos para resolverlos. En su mayoría, los trabajos existentes tratan 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 de forma simultánea. Si bien la comunidad científica ha propuesto varios enfoques de solución para atacar la optimización multi-objetivo, muchos de estos enfoques no se han aplicado aún en la ISBB. Uno de estos enfoques es el llamado “basado en preferencias”, el cual permite capturar e incorporar las preferencias entre los objetivos del tomador de decisiones, de manera de poder restringir el frente Pareto-óptimo a una zona de interés específica, con el objetivo de facilitar 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

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. 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.

Cabe 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 descripto 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 descriptos 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 multiobjetivo 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 descripto 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. Reformulación de problemas de la IS como problemas de búsqueda

El primer paso para aplicar cualquier técnica exitosamente es formular el problema como uno de búsqueda/optimización. En este sentido, se trabaja junto a expertos de la industria del software en la especificación de problemas de la IS en este formato requerido, adaptando y extendiendo los supuestos y características de los problemas del estado del arte, con el objetivo de darle utilidad a las soluciones que se pueden obtener con los distintos algoritmos disponibles.

2.2. 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].

2.3. 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 descripto 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.

2.4. Mecanismos de captación de preferencias

Como ya se dijo, los enfoques de optimización multiobjetivo 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

Hasta el momento se ha resuelto satisfactoriamente el problema del próximo lanzamiento (*Next Release Problem*, NRP) mediante múltiples algoritmos evolutivos como NSGA-II, IBEA y una versión novedosa de PSO denominada FMOPSO [12], desarrollada por los autores, escrita en C++. Además se construyó un prototipo en Python para obtener de forma exacta el frente de Pareto de un problema NRP bi-objetivo, utilizando GLPK (GNU Linear Programming Kit).

Nos es necesario además desarrollar 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 busca 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.

Lograr un primer prototipo de interfaz usable para usuarios no expertos en las técnicas de búsqueda heurística, es también un objetivo esperado. Este prototipo debe incorporar la posibilidad de realizar la carga de los datos necesarios de los problemas y la inducción de las preferencias de los usuarios.

Finalmente, se espera contar al final del proyecto con el conocimiento necesario para generar un *ranking* de algoritmos según distintos atributos al ser aplicados a distintos problemas de la ISBB.

4. FORMACIÓN DE RECURSOS HUMANOS

La investigación presentada constituye las líneas fundacionales de un nuevo grupo de investigación dentro de la UTN-FRCU, el GIICIS. Dos investigadores se encuentran realizando su tesis de doctorado. Un investigador está desarrollando su trabajo de especialización. Además participan en el proyecto dos becarios alumnos de la carrera Ingeniería en Sistemas de Información que inician su formación en la investigación, y dos alumnos realizan su práctica supervisada en el contexto de esta investigación.

5. BIBLIOGRAFÍA

[1] M. Harman and B. F. Jones, "Search-based software engineering," *Inf. Softw. Technol.*, vol. 43, no. 14, pp. 833–839, 2001.

[2] A. S. Sayyad, T. Menzies, and 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 and M. A. Ismail, "Research problems in Search-Based Software Engineering for many-objective optimization," in *2017 International Conference on Innovations in Electrical Engineering and Computational Technologies (ICIEECT)*, 2017, pp. 1–6.

[4] M. W. Mkaouer, M. Kessentini, S. Bechikh, and D. R. Tauritz, "Preference-based multi-objective software modelling," in *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, and Y. Zhang, "Search Based Software Engineering: Trends, Techniques and Applications," *ACM Comput. Surv.*, vol. 45, no. 1, p. 11, 2012.

[6] M. Harman, S. A. Mansouri, and 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, and 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, and 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 and G. Schweickardt, "Análisis de las habilidades de Metaheurísticas X-PSO Multiobjetivo 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.*, vol. 24, no. 40, pp. 23–42, 2016.

[11] M. Clerc, *Particle Swarm Optimization*. John Wiley & Sons, 2006.

[12] C. Casanova *et al.*, "Aproximación del Frente Pareto-Óptimo de un Problema NRP Bi-Objetivo mediante un Algoritmo basado en Enjambres de Partículas," in *Anales VI Seminario Argentina-Brasil de Tecnologías de la Información y la Comunicación*, 2018.

Patrones de Usabilidad Embebidos en la Construcción del Modelo Conceptual en Entornos Web

Juan Carlos Moreno[1], Marcelo Martín Marciszack[1], María Alejandra Paz Menvielle[1], Claudia Castro[1], Andrea Delgado[1], Silvio Serra[1], Juan Pablo Fernández Taurant[1], Ezequiel E. Fernández[1], Santiago Jose Solis Luna[1]
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, sant.solis.luna}@gmail.com [1]

RESUMEN

Detectar problemas de usabilidad en forma temprana, es un aspecto importante, que preocupa a la industria del software. El presente trabajo propone llevar a cabo un análisis de distintos tipos de patrones relacionados con criterios de usabilidad, para proponer una nueva metodología que brinde soluciones a distintos tipos de problemas de Usabilidad, no contemplados en las metodologías existentes de diseño web. El análisis se iniciará a partir de los Modelos de Procesos de Negocios, los cuales mediante transformaciones ayudarán a obtener los modelos conceptuales del sistema de información. En dichos modelos se introducirán patrones, que contribuirán a incorporar criterios de usabilidad en etapas tempranas del ciclo de vida del software. A partir del modelo obtenido anteriormente, se creará un conjunto de modelos abstractos que se emplearán para la validación de los modelos conceptuales diseñados. Esto permitirá obtener cierto grado de trazabilidad del esquema conceptual diseñado, ante posibles modificaciones o cambios de las especificaciones correspondientes, antes de que se inicie la codificación del mismo.

Palabras clave: Modelos de Negocios, Patrones de Usabilidad, MDA, Modelo Conceptual.

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). El presente trabajo empleará un método de investigación mixta, donde las unidades de observación estarán conformadas por expertos en desarrollo, y estudiantes universitarios en proceso de formación académica especializados en software. Se emplearán Revisiones sistemáticas de la literatura propuesta por Kitchenham et al. [6] y Genero et al. [7].

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 [3]. La calidad de las aplicaciones web se mide muchas veces basándose en el sentido común de los desarrolladores [4].

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 [5], 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 [6]. A su vez, la usabilidad es descompuesta en subatributos, haciendo que algunos atributos sean más tangibles y se puedan medir[7].

La norma ISO 25000 (Square) [8] 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 [9], [10]. 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) [11], 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 Arquitecture (MDA) (Miller y Mukerji 2003) [12]. 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 [13] [14], UWE [15] [16], OO-Method [17], OOH[18], OOWS[19], WebML[20].

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 sería 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: analizar un conjunto de patrones de diseño, con el fin de adaptarlos para resolver problemas de usabilidad en forma temprana en los modelos conceptuales.

En función de lo anterior, se plantearon las siguientes metas:

- Identificar patrones de diseño en los que se pueda incluir criterios de usabilidad en una etapa temprana de desarrollo en un entorno web.
- Proponer atributos y subatributos de usabilidad pueden incorporarse a los patrones de diseño identificados.
- Analizar el desempeño de los patrones generados a partir de la inclusión de los atributos de usabilidad.
- Evaluar las fortalezas y debilidades que pueda tener la modificación de los patrones de Diseño al incluir patrones de usabilidad.Identificar patrones de diseño en los que se pueda incluir criterios de usabilidad en una etapa temprana de desarrollo en un entorno web.
- Proponer atributos y subatributos de usabilidad pueden incorporarse a los patrones de diseño identificados.
- Analizar el desempeño de los patrones generados a partir de la inclusión de los atributos de usabilidad.
- Evaluar las fortalezas y debilidades que pueda tener la modificación de los patrones de Diseño al incluir patrones de usabilidad.
- Explorar la inclusión o uso de los patrones en etapas tempranas del proceso de desarrollo de software en entorno web.
- Caracterizar el uso de los patrones generados en relación con otras metodologías que aborden problemáticas similares.
- Explorar la inclusión o uso de los patrones en etapas tempranas del proceso de desarrollo de software en entorno web.
- Caracterizar el uso de los patrones generados en relación con otras metodologías que aborden problemáticas similares.

2. LÍNEAS 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:

Proponer 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 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.
- Definir un conjunto de patrones que permitan la inclusión de aspectos de usabilidad de forma temprana en el entorno de desarrollo web.

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". 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 Leóne.

Y finalmente se firmó un convenio de colaboración recíproca entre el Grupo de Investigación SI1-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] Kitchenham, B. and Charters, S. (2007). Guidelines for performing Systematic Literature Reviews in Software Engineering. Versión 2.3 EBSE-2007-01.

[2] Genero Bocco M., Cruz-Lemus J. y Piattini Velthuis M. (2014). Métodos de investigación en ingeniería del software. Editorial Ra-Ma, 2014. ISBN 978-84-9964-507-0.

[3] Pressman R., "What a tangled Wed we weave," *IEEE Software*, 2000.

[4] Abrahao S., Condori-Fernandez N., Olsina L., and Pastor O., "Defining and validating metrics for navigational models," Australia, 2003.

[5] Norma ISO/IEC ISO9126-1, "Software Engineering -Product Quality - Part 1," 2001.

[6] Nigel Bevan, "Quality and usability: A new framework," *Achieving software product quality*, 1997.

[7] Mario G. Piattini, Felix O. Garcia, and Ismael Caballero, "Calidad de Sistemas Informáticos," México, ISBN 978-970-15-1267-8, 2007.

[8] ISO/IEC 25000, Software Engineering - Software Product Quality Requirements and Evaluation (SQuaRE).

[9] L Bass and B John, "Linking usability to software arquitecture patterns through general scenarios," *The journal of systems and software*, no. 66, pp. 187 - 197, 2003.

[10] Eelke Folmer and Jan Bosh, "Architecting for usability: A survey. ," *Journal of Systems and Software*, pp. 61 - 78, 2004.

[11] Stephen J Mellor, Kendall Scott, Axel Uhl, and Dirk Weise, *Model-Driven Architecture*. Berlin / Heidelberg: Springer, 2002.

[12] (2003) MDA_Guide_Version1-0. [Online]. http://www.omg.org/mda/mda_files/MDA_Guide_Version1-0.pdf

[13] 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.

[14] S. Daniel, P. Rita de Almeida, and M. Isbela, "OOHDM-Web: an environment for implementation of hypermendia applications in the WWW," in *SIGWEB News l.8, 2*, 1999, pp. 18-34.

[15] Nora Koch and Martin Wirsing, *Software Engineering for Adaptative Hypermedia Applications*. München, Germany: Ludwig-Maximilians University of Munich, 2000.

[16] 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.

[17] 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.

[18] J. Gómez and C. Cachero, "OO-H Method: extending UML to model web interfaces.," in *In information Modeling For internet Applications,*. Hershey, PA.: Ed. IGI Publishing, , 2003, pp. 144-173.

[19] P.V., Albert M., and Pastor O. Fons J.,.: LNCS. Springer, 2003, vol. 2813, pp. 232-245.

[20] 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.

Procesos y Prácticas Ágiles de Desarrollo de Software

Nicolás Paez, Diego Fontdevila, Alejandro Oliveros
Departamento de Ciencia y Tecnología
Universidad Nacional de Tres de Febrero
nicopaez@computer.org, dfontdevila@untref.edu.ar, aoliveros@untref.edu.ar

RESUMEN

El presente artículo describe una línea de investigación sobre Procesos y Prácticas Ágiles de Desarrollo de Software llevada adelante en la Universidad Nacional de Tres de Febrero. Los temas abordados en esta línea se enfocan en el estudio de la adopción y la enseñanza de métodos ágiles de desarrollo de software. En este contexto se han desarrollado tres iniciativas en las cuales se ha trabajado realizando, principalmente, encuestas y reportes de experiencia. Los resultados obtenidos están en concordancia con lo esperado y también con algunos otros trabajos realizados a nivel global. En trabajos futuros se realizarán más encuestas y se comenzará a trabajar en casos de estudio para poder tener un abordaje de mayor profundidad.

Palabras clave: Métodos Ágiles, Ingeniería de Software, Prácticas de desarrollo

CONTEXTO

Las líneas de trabajo aquí descritas están enfocadas en "Procesos y Prácticas Ágiles de Desarrollo de Software" y se encuentran enmarcadas en el proyecto de investigación "Procesos de Desarrollo de Software" perteneciente a la Universidad Nacional de Tres de Febrero. El proyecto marco incluye también otras líneas de investigación que exceden lo incluído en el presente artículo. El grupo de trabajo está conformado por docentes e investigadores de la carrera de Ingeniería en Computación. El trabajo aquí presentado comenzó a desarrollarse en 2016 y el plan de trabajo actual se extiende hasta fines de 2019.

1. INTRODUCCIÓN

Los métodos ágiles de desarrollo software han crecido enormemente en popularidad desde la firma del manifiesto ágil en 2001 [1] y se han convertido en *mainstream*. Esto queda evidenciado por la gran oferta de conferencias, libros y cursos disponibles en la actualidad. Tanto la industria como la academia han incorporado los métodos ágiles.

Estos métodos tienen una visión holística del desarrollo de software que incluye tanto cuestiones de gestión como cuestiones de índole técnica. Los enfoques ágiles pueden ser al mismo tiempo simples pero difíciles de poner en práctica. No es una cuestión de solo entender los principios y fundamentos de estos métodos sino también de aprender y aplicar efectivamente las prácticas asociadas. Como suele ocurrir con las ideas exitosas, muchos han intentado utilizar métodos ágiles, pero no todos ellos han logrado una adopción completamente exitosa. Algunos referentes del movimiento ágil se han expresado sobre implementaciones fallidas métodos ágiles [2, 3].

Resulta entonces relevante entender las características de la adopción de las prácticas ágiles para poder en base a ello mejorar los mecanismos de adopción. La gran mayoría de los estudios sobre el uso de las prácticas ágiles en el desarrollo de software están enfocados en los beneficios y los desafíos que observan los practicantes. Algunos ejemplos de esto son [4, 5, 6, 7]. No hay estudios que analicen el uso de prácticas ágiles desde la perspectiva de

prácticas técnicas y prácticas de gestión. Tampoco hay estudios recientes sobre el grado de uso de prácticas ágiles en la región. Sin embargo, hay algunos estudios con cierto nivel de relación que vale la pena destacar. Ambler publicó en 2009 una encuesta [8] sobre el uso de prácticas ágiles, la cual a pesar de no ser un estudio formal, da una primera idea del nivel de adopción de las distintas prácticas. La empresa Version One realiza ya desde hace varios años una encuesta anual a nivel mundial sobre el uso de Agile [9]. En 2013 en Brasil se realizó un estudio [10] sobre la evolución de los métodos ágiles, el mismo abarca tanto cuestiones de industria como de academia, pero sin entrar en detalle sobre el uso de prácticas. Diebold publicó una cartografía sistemática sobre Prácticas en la Práctica [11]. Kropp estudió el uso de prácticas ágiles en Suiza caracterizando prácticas técnicas y de colaboración [12]. Mientras que Pantiuchina estudio el uso de prácticas ágiles en *startups* [7] haciendo la diferenciación de prácticas orientadas a la calidad y prácticas orientadas a la velocidad.

En los últimos años, algunos autores [13, 14] han comenzado a hablar de métodos híbridos de desarrollo de software haciendo referencia a estrategias de desarrollo, cada vez más habituales en la industria, que incorporan prácticas de distintas metodologías sean estas de origen ágil o no.

Todo lo descripto en los párrafos anteriores da pie a nuestra motivación de estudiar el uso de métodos y prácticas ágiles enfocándonos particularmente en América Latina y considerando la relación de prácticas técnicas y prácticas de gestión.

2. LÍNEAS DE INVESTIGACIÓN

Dentro de nuestro proyecto hemos establecido tres línea de trabajo.

Nuestra primera y principal línea de trabajo está relacionada al uso de prácticas ágiles en América Latina. En este sentido pretendemos entender el estado del arte en la adopción de las prácticas ágiles de cara a poder dar respuesta a las siguientes preguntas:

- ¿Cuál es el grado de adopción de las distintas prácticas ágiles en los proyectos de desarrollo de software realizados en la región?
- ¿Cuál es la relación, en términos de adopción de las prácticas de gestión versus las prácticas técnicas?
- ¿En qué medida las prácticas reales se ajustan a la conceptualización establecida de dichas prácticas en la comunidad ágil?
- ¿Los resultados obtenidos en la Argentina son consistentes con el estado de la utilización de las prácticas en la región?

Nuestra segunda línea de trabajo es en el contexto de una iniciativa internacional denominada HELENA: Hybrid dEveLopmENt Approaches in software systems development [13]. Esta línea está enfocada en el estudio de métodos híbridos de desarrollo de software y el objetivo es entender el grado de uso de métodos híbridos y cómo es la composición de esa mezcla híbrida.

Nuestra tercera línea de trabajo apunta a profundizar sobre las prácticas ágiles y entender cómo es la enseñanza de métodos ágiles en América Latina. Aquí queremos entender en qué medida y de qué forma la academia está cubriendo la formación en métodos ágiles.

Desde el punto de vista metodológico, todos nuestros trabajos hasta el momento se han basado en la realización de encuestas. En la mayoría de los casos las encuestas las hemos realizado en forma presencial en el contexto de conferencias. A partir de los hallazgos realizados en las encuestas, planeamos realizar casos de estudio basados en entrevistas con referentes de las correspondientes temáticas.

3. RESULTADOS OBTENIDOS/ESPERADOS

Los resultados de este trabajo de investigación han sido publicados incrementalmente en distintas conferencias.

En lo referente al uso de prácticas ágiles, realizamos 3 encuestas y los hallazgos de cada una de ellas fueron publicados en un artículo de conferencia y sirvieron como punto de partida de la siguiente encuesta [15, 16, 17]. En forma resumida los hallazgos destacados de estas tres publicaciones fueron:

- Una marcada diferencia entre el uso de prácticas ágiles de gestión y prácticas ágiles de carácter técnico, siendo las prácticas de gestión mucho más utilizadas que las prácticas técnicas
- Un defasaje entre la definición de ciertas prácticas y la puesta en práctica de las mismas por parte de los practicantes.
- Una clara correlación entre la cantidad de prácticas utilizadas y los años de experiencia de la organización en el uso de métodos ágiles.
- Los resultados obtenidos en nuestro estudio en América Latina son consistentes con resultados de otros estudios realizados a nivel mundial.

En lo referente a los métodos híbridos, parte de nuestros resultados fueron publicados en [18] y otra parte está aún pendiente de publicación.

Finalmente, hemos publicado otros dos artículos que son colaterales a las dos líneas antes mencionadas: un reporte de experiencia sobre la enseñanza de ingeniería de software utilizando prácticas ágiles [19] y un reporte de experiencia sobre la práctica de versionado en contexto de DevOps [20].

Durante 2018 comenzamos a trabajar en una nueva iniciativa con el objetivo poder generar nuevas hipótesis de trabajo. Esta iniciativa consistió en la realización de talleres sobre algunas prácticas ágiles como Test-Driven Development, Pair-Programming e Integración Continua. Estos talleres fueron realizados en el contexto de dos conferencias: el Congreso Nacional de Ingeniería Informática y Sistemas de Información (CONAIISI 2018) y las Jornadas de Ingeniería de Software del Uruguay (JIS2018).

Actualmente estamos trabajando en dos cuestiones: entender la enseñanza de prácticas ágiles en las carreras de informática y en el análisis y publicación de segunda parte de los resultados de nuestros estudio de métodos híbridos.

A futuro, tenemos planeado profundizar en el estudio del uso de prácticas ágiles, realizando casos de estudio, y también profundizar en el estudio de la enseñanza de prácticas ágiles, a partir de la realización de una encuesta a docentes del área de Ingeniería de Software.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de proyecto está conformado por un director, un co-director (que se encuentra haciendo su doctorado) y un investigador en formación. En la primera etapa del proyecto (2016-2017) también fue parte del equipo un alumno de grado de la carrera de Ingeniería en Computación de UNTreF.

Adicionalmente, en los talleres de Test-Driven Development mencionados en la sección precedente, participaron un total de 27 personas.

5. REFERENCIAS

[1] K. Beck et al, “Manifesto for Agile Software Development.”. Disponible en línea: http://agilemanifesto.org/. Accedido: 01-03-2019.

[2] M. Fowler, “Flaccid Scrum”. Disponible en línea: https://martinfowler.com/bliki/FlaccidScrum.html. Accedido: 01-03-2019

[3] J. Kerievsky, “Stop Calling them Technical Practices”. Disponible en línea: https://www.linkedin.com/pulse/stop-calling-them-technical-practices-joshua-kerievsky/?published=t. Accedido: 01-03-2019

[4] T. Dybå & T. Dingsøyr, “Empirical studies of agile software development: A systematic

review". *Information and Software Technology, 50(9–10), 833–859, 2008* http://doi.org/10.1016/j.infsof.2008.01.006

[5] A. Aldahmash, A. M. Gravell, & Y. Howard, "A review on the critical success factors of agile software development". In: Stolfa, J., Stolfa, S., O'Connor, R.V., Messnarz, R. (eds.) EuroSPI 2017. CCIS, vol. 748, pp. 504–512. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-64218-5_41

[6] T. Chowy & D.B. Cao, "A survey study of critical success factors in agile software projects". *Journal of Systems and Software, 81(6), 961–971.* http://doi.org/10.1016/j.jss.2007.08.020

[7] J. Pantiuchina, M. Mondini, D. Khanna, X. Wang, & P. Abrahamsson, "Are Software Startups Applying AgilePractices? The State of the Practice from a Large Survey". *InAgile Processes in Software Engineering and Extreme Programming (pp. 167-183).* Springer, Cham. https://doi.org/10.1007/978-3-319-57633-6_11

[8] S. Ambler, "Agile Practices Survey Results", July 2009. Disponible en línea: http://www.ambysoft.com/surveys/practices2009.html. Accedido: 01-03-2019

[9] Version One, " State of Agile Development Survey", Disponible en línea: https://stateofagile.versionone.com/. Accedido: 01-03-2019

[10] O. Melo, C.V. Santos, & E. Katayama,"The Evolution of Agile Software Development in Brazil." *Journal of the Brazilian Computer Society* 19, no. November 2013 (n.d.): 523–552.

[11] P. Diebold & M. Dahlem, "Agile practices in practice: a mapping study", *In Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, p. 30. ACM, 2014.

[12] M. Kropp, A. Meier & R. Biddle, "Agile practices, collaboration and experience". *In: Abrahamsson, P., Jedlitschka, A., Nguyen Duc, A., Felderer, M., Amasaki, S., Mikkonen, T. (eds.) PROFES 2016. LNCS, vol. 10027, pp. 416–431. Springer, Cham* (2016). https://doi.org/10.1007/978-3-319-49094-6_28

[13] M. Kuhrmann et al, "On the use of hybrid development approaches in software and systems development: construction and test of the HELENA survey" *Proceedings of the Annual Special Interest Group Meeting Projektmanagement und Vorgehensmodelle (PVM)(Lecture Notes in Informatics (LNI)). Vol. 263,* (2016)

[14] G. Theocharis et al. "Is Water-Scrum-Fall Reality? On the Use of Agile and Traditional Development Practices". *In Product-Focused Software Process Improvement: 16th International Conference, PROFES 2015, Bolzano, Italy, December 2-4, 2015, Proceedings. ed. / Pekka Abrahamsson; Luis Corral; Markku Oivo; Barbara Russo. Springer, 2015. p. 149-166*

[15] N. Paez, D. Fontdevila, & A. Oliveros, "Characterizing technical and organizational practices in the Agile Community" *In: Proceedings of CONAIISI, Salta, Argentina,* (2016)

[16] N. Paez, F. Gainey, A. Oliveros & D. Fontdevila, "An empirical study on the usage of technical and organizational practices in the Agile Community". *In: Proceedings of CONAIISI, Santa Fe, Argentina* (2017)

[17] N. Paez, D. Fontdevila, F. Gainey & A. Oliveros, "Technical and Organizational Agile Practices: A Latin-American Survey" *In: Garbajosa J., Wang X., Aguiar A. (eds) Agile Processes in Software Engineering and Extreme Programming. XP 2018. Lecture Notes in Business Information Processing, vol 314. Springer, Cham*

[18] N. Paez, D. Fontdevila, & A. Oliveros, "HELENA study: Initial observations of software development practices in Argentina". *In: Felderer, M., Méndez Fernández, D., Turhan, B., Kalinowski, M.,*

Sarro, F., Winkler, D. (eds.) PROFES 2017. LNCS, vol. 10611, pp. 443–449. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-69926-4_34

[19] N. Paez, "A Flipped Classroom Experience Teaching Software Engineering", *2017 IEEE/ACM 1st International Workshop on Software Engineering Curricula for Millennials (SECM), Buenos Aires*, 2017, pp. 16-20.doi: 10.1109/SECM.2017.6

[20] N. Paez, "Versioning Strategy for DevOps Implementations", *2018 Congreso Argentino de Ciencias de la Informática y Desarrollos de Investigación (CACIDI), Buenos Aires*, 2018, pp. 1-6.doi: 10.1109/CACIDI.2018.8584362

Reuso de Servicios basado en Análisis de Compatibilidad y Complejidad

Andres Flores, Martin Garriga, Alejandra Cechich,
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, martin.garriga]@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 GI 11220170100951CO: "Construcción de Líneas de Productos Software guiada por Estándares de Dominio". Financiado por CONICET. (2018-2020).
- PICT-2017-1725: "Extensión de Metamodelos de Aplicaciones Orientadas a Servicios para Descripción, evaluación y despliegue de Servicios en la Nube". Financiado por ANPCyT. (12/2018-12/2020).
  - Investigaciones conjuntas con ISISTAN-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 [FCGMRAC18] hemos enumerado una serie de contribuciones anteriores. Durante el año 2018, 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) [GRLFMCZ18].

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 y 1 investigador asistente CONICET), 2 doctorandos y 1 maestrando entre los miembros del proyecto. Dirección de Tesis de Grado durante 2018: 4 tesis.

Referencias

[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.

[GRLFMCZ18] Garriga, M., De Renzis, A., Lizarralde, I., Flores, A., Mateos, C., Cechich, A., Zunino, A. (2016). *A Structural-Semantic Web Service Selection Approach to Improve Retrievability of Web Services*. Information Systems Frontiers, 20(6): 1319-1344 (**Indexed SCI IF JCR2017: 3.232**).

[FCGMRAC18] Flores, A., Cechich, A., Garriga, M., Moyano, M., De Renzis, A., Anabalon, D., Corgatelli, F. (2018). *Evaluación de Compatibilidad y Complejidad para el Reuso de Servicios*. WICC'18, XX Workshop de Investigadores en Ciencias de la Computación. Abril, Corrientes.

[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.

[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.

Reuso orientado a dominios: Hacia un enfoque integral de familias de procesos y productos

Agustina Buccella, Juan Luzuriaga, Alejandra Cechich,
Matias Pol'la, Maximiliano Arias, Marcos Cruz
Rodolfo Martínez, Rafaela Mazalu
GIISCO Research Group
Departamento de Ciencias de la Computación
Universidad Nacional del Comahue
Neuquen, Argentina
agustina.buccella@fi.uncoma.edu.ar

1. Resumen

La gestión de procesos de negocio (Business Process Management -BPM) es un conjunto de métodos, herramientas y tecnologías utilizadas para diseñar, representar, analizar y controlar los procesos operacionales de las organizaciones. Cuando trabajamos con organizaciones grandes o medianas, se debe considerar que los procesos no pueden ser definidos en forma estática y fija, sino que deben poder adaptarse o personalizarse a los cambios y contextos en donde se ejecutan los mismos. Para lograr esta adaptación, se aplican mecanismos similares a los utilizados en el área de las Líneas de Productos de Software (LPS) en donde la definición, modelado, implementación e instanciación de la variabilidad son actividades fuertemente investigadas en la actualidad.

Nuestro proyecto actual se basa principalmente en el área de las Líneas de Productos de Software proponiendo enfoques, métodos, técnicas y recursos para lograr mejoras dentro de un proceso integral de desarrollo. Dicho proceso ha sido aplicado dentro de una jerarquía de dominios que comparten generalidades y particularidades en forma de servicios comunes y variables. En base a esta línea de investigación proponemos extender y crear los mecanismos y recursos necesarios para la definición de modelos de procesos de negocio con variabilidad dentro de los dominios ya analizados de manera de fomentar el reuso en ambos niveles. En particular, en una primera etapa nos centraremos en el análisis de los modelos de negocio variables en cuanto a las anomalías, redundancias e inconsistencias que se pueden encontrar.

Palabras Clave: Gestión de Procesos de Negocio - Líneas de Producto de Software - Análisis de Variabilidad - Dominios - Reuso

2. Contexto

La línea presentada se inserta en el contexto del *Programa: Desarrollo de Software Basado en Reuso - Parte II (04/F009)*. Directora: Dra. Alejandra Cechich, y *SubProyecto: Reuso Orientado a Dominios - Parte II*. Incluído dentro del Programa. Directora: Dra. Agustina Buccella, Codirector: Mg. Juan Manuel Luzuriaga.

3. Introducción

La gestión de procesos de negocio (Business Process Management -BPM) [30, 31] es una área de la ingeniería de software ampliamente investigada en la actualidad. Existen muchas definiciones en la literatura del significado de lo que ésta área significa, por ejemplo en [30] se define como: *Asistir a los procesos de negocios utilizando métodos, técnicas y software para diseñar, promulgar, controlar y analizar procesos operacionales que involucran a personas, organizaciones, aplicaciones, documentos y otras fuentes de información*. A su vez, cuando se realiza un proceso de BPM sobre organizaciones medianas y grandes, es muy común que los procesos de negocios posean variantes basadas en requisitos comerciales específicos, por ejemplo, diferentes procesos de venta según el tipo de productos o el método de pago, o diferentes pasos de rendición de cuentas según el país en el que se ejecuta el proceso [12]. La definición de estas variantes en etapas tempranas del desarrollo de un BPM brindarán flexibilidad y adaptabilidad dentro de los

procesos para contemplar las diferentes configuraciones que pueden existir según el contexto de implementación de un modelo de proceso de negocio en particular. Lograr la flexibilidad para adaptarse a los contextos dinámicos se denomina *variabilidad de procesos de negocios*, el cual es un campo emergente en BPM con muchas de sus propuestas inspiradas en el área de las Líneas de Productos de Software (LPS) [11, 24] pero aplicadas a manejar la variabilidad en procesos. Así, la gestión de la variabilidad es, tanto para LPS como para BPM, un conjunto de tareas o actividades dedicadas a proporcionar flexibilidad y un alto nivel de reutilización durante el desarrollo de software o definición de procesos de negocio, respectivamente. Dentro del enfoque de las Líneas de Productos de Software, las actividades relacionadas con la gestión de variabilidad son las encargadas de proveer la flexibilidad necesaria para desarrollar un conjunto de aplicaciones similares basadas en un rango manejable de funcionalidades variables según las necesidades de los usuarios expertos. En este sentido, dentro de BPM la gestión de la variabilidad involucra también un conjunto de actividades que permitan la definición, modelado, implementación y derivación de la variabilidad desde las fases tempranas del ciclo de vida de BPM. De esta forma se permite la creación de familias de procesos, es decir, un conjunto de procesos de negocios que se basan en un proceso base (también conocido como: proceso personalizable) y un conjunto de puntos de variación que son elementos del proceso base que se pueden personalizar según el contexto. Así, la variabilidad en los procesos de desarrollo de software permite la personalización de actividades, artefactos, roles y otros elementos para proyectos específicos, en lo que se denomina una variante de proceso.

Existen varios trabajos en la literatura que proponen metodologías para incluir variabilidad en las etapas tempranas del ciclo de vida de BPM. En general, los lenguajes de modelado de procesos, como por ejemplo, el estándar BPMN 2.0[1] (Modelo de Proceso de Negocio y Notación), no admite explícitamente la especificación de familias de procesos, es decir, no permite definir procesos personalizables con un conjunto de puntos de variación [27]. Sin embargo, hay varias propuestas recientes que extienden o crean nuevos lenguajes y notaciones que incluyen variabilidad [4, 17, 27, 28, 29]. Por ejemplo en [27, 29] los autores presentan una revisión literaria y un survey sobre variabilidad en procesos de negocio.

A su vez hay propuestas de enfoques de variabilidad en procesos basados en dominios específicos como es el caso de procesos software, es decir procesos de negocio involucrados en el desarrollo de software. Dentro de este dominio un lenguaje estándar de alto nivel es SPEM (Software Process Engineering Metamodel)[2]. A pesar de que el estándar define elementos específicos para modelar la variabilidad, definiendo varios tipos de variabilidad entre dos elementos relacionados, posee limitaciones desde la perspectiva de un proceso software [13]. En este sentido han surgido varias extensiones en la literatura para solucionar dichas limitaciones y proponer mejoras en varios aspectos de la gestión de la variabilidad [2, 13, 18, 19, 22].

Por otro lado, otro aspecto a ser analizado dentro de la variabilidad en procesos de negocio son aquellas propuestas que analizan específicamente los aspectos centrados en el campo del *análisis automático de la variabilidad* [5]. Dicho campo, que surge del área de las Líneas de Producto de Software, se centra específicamente en validar modelos de variabilidad de acuerdo con un conjunto predefinido de escenarios de validación [5, 16, 26]. En la literatura, existen muchos trabajos sobre el análisis (automatizado) de modelos de variabilidad en el campo de las LPSs. Por ejemplo, en el trabajo de Benavides etal. [5] se presenta un proceso general que define un conjunto de tareas para evaluar modelos de variabilidad, específicamente modelos de características (feature models - FMs). En la Figura 1 podemos ver este proceso, en la cual se identifican cinco componentes principales, definiendo un proceso general que sienta las bases de la actividad de la verificación automática de modelos de variabilidad.

El primer componente es el *modelo de variabilidad* que puede definirse siguiendo cualquiera de los diferentes enfoques de modelado. El segundo componente es el *traductor*, que realiza el proceso de transformación entre el modelo y la representación formal teniendo en cuenta el enfoque seleccionado. El tercer componente es el *modelo formal* o representación formal, que incluye la variabilidad en términos lógicos. El cuarto componente es un *solver*, responsable de validar el modelo formal. Este componente también recibe el conjunto de escenarios o consultas de validación, que determinan los *resultados* (quinto componente) del proceso de análisis.

[1] `http://www.bpmn.org/`

[2] `https://www.omg.org/spec/SPEM/About-SPEM/`

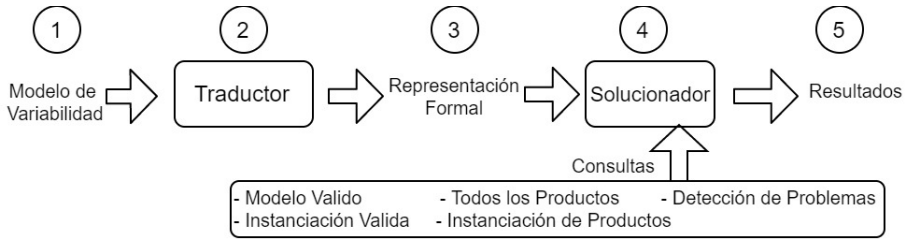


Figura 1. Proceso general para analisis automatico de modelos de variabilidad en LPS

En el área de variabilidad en procesos de negocio, hay también varios avances que aplican un proceso para el análisis/verificación de variabilidad en modelos de características aplicados a los procesos de negocio. Es decir, propuestas como [1, 3, 20] definen la variabilidad de los procesos de negocio mediante el uso de diagramas de características. En estos casos, se aplican varias de las técnicas ya conocidas en las LPSs para analizar variabilidad también para estos modelos. Luego, otras propuestas en la literatura que se centran específicamente en el análisis de los modelos de procesos y su variabilidad utilizando otros lenguajes de modelado pueden ser vistos en [14, 15, 21] de los cuales [14, 21] representan surveys o revisiones sistemáticas.

4. Líneas de Investigación y Desarrollo

En los últimos años hemos realizado varios avances en el modelado de LPS para el dominio geográfico general y en los subdominios de ecología marina y paleontología en particular. Dentro de estos avances hemos definido la base para la división o jerarquía de dominios en subdominos más manejables [7, 8] que mejoraron la complejidad inherente del dominio geográfico general. Para esto, se ha trabajado dentro de los subdominios de ecología marina y paleontología, en los cuales hemos desarrollado una taxonomía de servicios [8, 23] basada en los estándares geográficos definidos por el OGC[3] y la ISO[4].

Al mismo tiempo hemos definido una metodología de desarrollo [7, 9], la cual se basa en la construcción de artefactos de software para cada una de las actividades, comenzando con la taxonomía previamente mencionada. Son justamente estos artefactos los que guían a las actividades de manera de mejorar el desarrollo de sistemas en el dominio geográfico y en sus subdominios incluidos.

A su vez, dentro del campo del análisis de los modelos de variabilidad, hemos desarrollado dos herramientas basadas en diferentes representaciones lógicas y solucionadores [6, 25]. Ambos, toman modelos de variabilidad como entradas (construidos por nuestro propio enfoque, llamados modelo SeVaTax [7]) y hacen diferentes traducciones para ser evaluadas por diferentes solucionadores. En [25] se presenta una implementación en CNF (Conjuntive Normal Form) junto con SAT solver (SAT4j[5]); y en [6], presentamos otra implementación en DL (Description Logic) y usando DL Reasoners (RACER[6]). Sin embargo, estas composiciones de herramientas son solo dos posibles implementaciones que se pueden usar bajo requisitos específicos. Así, por último en [10] hemos propuesto un framework para combinar herramientas de análisis de variabilidad automáticas, que pueden adaptarse a los requisitos específicos de un caso de gestión de variabilidad.

5. Resultados Obtenidos/Esperados

El objetivo principal de la línea de investigación es *Desarrollar técnicas y herramientas que mejoren el desarrollo de software basado en reuso a nivel de dominios, favoreciendo el desarrollo de ambientes inteligentes que permitan crear LPSs según taxonomías de servicios*.

Dentro de este objetivo, se definen dos líneas de investigación específicas. En la primera, hemos obtenido resultados concretos y continuaremos trabajando dentro del área de Líneas de Productos de Software. Como se ha descripto los trabajos previos de la sección anterior, hemos avanzado en el desarrollo de artefactos de software y herramientas que sirven de soporte para las actividades dentro de la metodología de LPS definida (meta-modelos, taxonomías de servicio, etc.), hemos definido y aplicado nuevos métodos y técnicas para la creación de LPSs con soportes inteligentes dentro del dominio geográfico que contemplen las particularidades de los subdominios incluidos, y hemos realizado

[3] http://www.opengeospatial.org/
[4] http://www.isotc211.org/
[5] www.sat4j.org/
[6] https://www.ifis.uni-luebeck.de/~moeller/racer/index.html

formalizaciones de reglas y patrones para soportar el desarrollo asistido, de manera que sean lo suficientemente generales para ser aplicados en otros subdominios geográficos.

En la segunda línea derivada del objetivo principal, se comenzará a trabajar en la definición de los modelos de procesos de negocios con variabilidad bajo el mismo contexto con el que se definieron los dominios y subdominios de las LPSs. Se deberá analizar la forma en que la taxonomía de servicios puede ser utilizada para la definición de los procesos y la necesidad de crear otros artefactos de software o recursos útiles para incrementar el reuso asistiendo a la creación de otros modelos para otros subdominios. Al mismo tiempo, continuaremos trabajando en el campo del análisis de la variabilidad definiendo modelos de procesos de negocio con variabilidad y proponiendo la incorporación de mecanismos que permitan analizar automáticamente la variabilidad dentro de dichos modelos. De esta manera, el proceso general para el análisis automático de los modelos de variabilidad de las LPS (Figura 1) deberá ser adaptado en base a las consultas o escenarios de validación que se definan de forma de tener identificadas y formalizadas las anomalías inconsistencias y redundancias que se puedan encontrar en los modelos de procesos y que perjudiquen la correcta implementación de un BPM. Específicamente trabajaremos en el subdominio de paleontología.

De esta manera, ambas líneas contribuyen a la creación de ambientes inteligentes que permitan fomentar el reuso de artefactos de software en el ámbito de los productos y de los procesos.

6. Formación de Recursos Humanos

El proyecto reúne aproximadamente a 13 investigadores, entre los que se cuentan docentes y alumnos de UNComa, y colaboradores. A su vez, el proyecto cuenta actualmente con un doctor y un magister. Varios de los docentes-investigadores de GIISCo-UNComa han terminado o se encuentran próximos a terminar carreras de postgrado. Uno de ellos se encuentra finalizando su doctorado en el transcurso de este año. A su vez varios de los integrantes se encuentran finalizando sus tesis de grado. Por último, este año contaremos con la incorporación de 2 becarios EVC-CIN.

Referencias

[1] Mathieu Acher, Philippe Collet, Philippe Lahire, and Robert France. Managing variability in workflow with feature model composition operators. In *Proceedings of the 9th International Conference on Software Composition*, SC'10, pages 17–33, Berlin, Heidelberg, 2010. Springer-Verlag.

[2] J. A. H. Alegría and M. C. Bastarrica. Building software process lines with casper. In *2012 International Conference on Software and System Process (ICSSP)*, pages 170–179, June 2012.

[3] Mohsen Asadi, Bardia Mohabbati, Gerd Gröner, and Dragan Gasevic. Development and validation of customized process models. *Journal of Systems and Software*, 96:73 – 92, 2014.

[4] Clara Ayora, Victoria Torres, Barbara Weber, Manfred Reichert, and Vicente Pelechano. Vivace: A framework for the systematic evaluation of variability support in process-aware information systems. *Information and Software Technology*, 57:248 – 276, 2015.

[5] D. Benavides, S. Segura, and A. Ruiz-Cortés. Automated analysis of feature models 20 years later: A literature review. *Information Systems*, 35:615–636, September 2010.

[6] Germán Alejandro Braun, Matias Pol'la, Laura Andrea Cecchi, Agustina Buccella, Pablo R. Fillottrani, and Alejandra Cechich. A DL semantics for reasoning over ovm-based variability models. In *Proceedings of the 30th International Workshop on Description Logics, Montpellier, France, July 18-21, 2017.*, 2017.

[7] A. Buccella, A. Cechich, M. Arias, M. Pol'la, S. Doldan, and E. Morsan. Towards systematic software reuse of gis: Insights from a case study. *Computers & Geosciences*, 54(0):9 – 20, 2013.

[8] A. Buccella, A. Cechich, M. Pol'la, M. Arias, S. Doldan, and E. Morsan. Marine ecology service reuse through taxonomy-oriented SPL development. *Computers & Geosciences*, 73(0):108 – 121, 2014.

[9] A. Buccella, M. Pol'la, M. Arias, and A. Cechich. Taxonomy-based annotations for software product line development. In *17 Simposio Argentino de Ingeniería de Software (ASSE'16) en el marco de las 45 Jornadas Argentinas de Informática (JAIIO)*, Buenos Aires, Argentina, 2016.

[10] Agustina Buccella, Matias Pol'la, Esteban Ruiz de Galarreta, and Alejandra Cechich. Combining automatic variability analysis tools: An spl approach for building a framework for composition. In Osvaldo Gervasi, Beniamino Murgante, Sanjay Misra, Elena Stankova, Carmelo M. Torre, Ana Maria A.C. Rocha, David Taniar, Bernady O. Apduhan, Eufemia Tarantino, and Yeonseung Ryu, editors, *Computational Science and Its Applications – ICCSA 2018*, pages 435–451, Cham, 2018. Springer International Publishing.

[11] P. C. Clements and L. Northrop. *Software Product Lines : Practices and Patterns*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[12] A. Delgado and D. Calegari. Bpmn 2.0 based modeling and customization of variants in business process families. In *2017 XLIII Latin American Computer Conference (CLEI)*, pages 1–9, Sep. 2017.

[13] Andrea Delgado, Daniel Calegari, and Félix García. Modeling of software process families with automated generation of variants (S). In *The 30th International Conference on Software Engineering and Knowledge Engineering, Hotel Pullman, Redwood City, California, USA, July 1-3, 2018.*, pages 330–329, 2018.

[14] H Groefsema and Doina Bucur. A survey of formal business process verification from soundness to variability. *BMSD 2013 - Proceedings of the 3rd International Symposium on Business Modeling and Software Design*, pages 198–203, 01 2013.

[15] Gerd Gröner, Marko Bošković, Fernando Silva Parreiras, and Dragan Gašević. Modeling and validation of business process families. *Information Systems*, 38(5):709 – 726, 2013.

[16] 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*, GPCE 2016, pages 132–143, New York, NY, USA, 2016. ACM.

[17] Marcello La Rosa, Wil M. P. van der Aalst, Marlon Dumas, and Arthur H. M. ter Hofstede. Questionnaire-based variability modeling for system configuration. *Software & Systems Modeling*, 8(2):251–274, Apr 2009.

[18] T. Martinez-Ruiz, F. Garcia, M. Piattini, and J. Munch. Modelling software process variability: an empirical study. *IET Software*, 5(2):172–187, April 2011.

[19] Tomás Martínez-Ruiz, Félix García, and Mario Piattini. *Towards a SPEM v2.0 Extension to Define Process Lines Variability Mechanisms*, pages 115–130. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[20] Ildefonso Montero, Joaquin Pena, and Antonio Ruiz-Cortes. From feature models to business processes. In *Proceedings of the 2008 IEEE International Conference on Services Computing - Volume 2*, SCC '08, pages 605–608, Washington, DC, USA, 2008. IEEE Computer Society.

[21] Shoichi Morimoto. A survey of formal verification for business process modeling. In Marian Bubak, Geert Dick van Albada, Jack Dongarra, and Peter M. A. Sloot, editors, *Computational Science – ICCS 2008*, pages 514–522, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

[22] Edson A. Oliveira Junior, Maicon G. Pazin, Itana M. S. Gimenes, Uirá Kulesza, and Fellipe A. Aleixo. Smartyspem: A spem-based approach for variability management in software process lines. In Jens Heidrich, Markku Oivo, Andreas Jedlitschka, and Maria Teresa Baldassarre, editors, *Product-Focused Software Process Improvement*, pages 169–183, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[23] Fiorella Pesce, Sofia Caballero, Agustina Buccella, and Alejandra Cechich. Reusing a geographic software product line platform: A case study in the paleontological sub-domain. In Armando Eduardo De Giusti, editor, *Computer Science – CACIC 2017*, pages 145–154, Cham, 2018. Springer International Publishing.

[24] 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.

[25] Matias Pol'la, Agustina Buccella, and Alejandra Cechich. Automated analysis of variability models: The sevatax process. In *Computational Science and Its Applications - ICCSA 2018 - 18th International Conference, Melbourne, VIC, Australia, July 2-5, 2018, Proceedings, Part IV*, pages 365–381, 2018.

[26] Fabricia Roos-Frantz, José A Galindo, David Benavides, Antonio Ruiz Cortés, and J Garcıa-Galán. Automated analysis of diverse variability models with tool support. *Jornadas de Ingenierıa del Software y de Bases de Datos (JISBD 2014), Cádiz. Spain*, page 160, 2014.

[27] Marcello La Rosa, Wil M. P. Van Der Aalst, Marlon Dumas, and Fredrik P. Milani. Business process variability modeling: A survey. *ACM Comput. Surv.*, 50(1):2:1–2:45, March 2017.

[28] M. Rosemann and W.M.P. van der Aalst. A configurable reference modelling language. *Information Systems*, 32(1):1 – 23, 2007.

[29] George Valenca, Carina Alves, Vander Alves, and Nan Niu. A systematic mapping study on business process variability. *International Journal of Computer Science and Information Technology*, 5(1):1–21, feb 2013.

[30] Wil M. P. van der Aalst, Arthur H. M. ter Hofstede, and Mathias Weske. Business process management: A survey. In Wil M. P. van der Aalst and Mathias Weske, editors, *Business Process Management*, pages 1–12, Berlin, Heidelberg, 2003. Springer Berlin Heidelberg.

[31] Mathias Weske. *Business Process Management: Concepts, Languages, Architectures*. Springer-Verlag, Berlin, Heidelberg, 2007.

Rúbrica para la Evaluación de Equipos de Desarrollo de Software

Marcelo Estayno, Fabiana Grinsztajn
Universidad Nacional de San Martín

mestayno@unsam.edu.ar,fabianagrin@gmail.com

RESUMEN

La investigación aborda el estudio de las metodologías que en la actualidad son utilizadas en el desarrollo del software, se focalizó en dos localidades maduras para este tipo de desarrollo industrial, en la República Argentina (Ciudad de Buenos Aires y Ciudad de Córdoba).

En una primera etapa se analizó la perspectiva de desarrolladores, identificando las competencias clave valoradas y analizando el papel que desempeña la comunicación y la trama vincular entre los stakeholders (Kuechler, 1988).

La finalidad ha sido el diseño de un marco de trabajo y un conjunto de recomendaciones para buenas prácticas, a partir de la evaluación que surge de la investigación.

Para hacer el estudio, se analizaron antecedentes vinculados con la temática y se administró una encuesta y grupos focales.

A patir del analisis de los resultados obtenidos, se construyó una herramienta de evaluación del funcionamiento de los equipos de desarrollo de software mediante el uso de una rúbrica, inserta en la ceremonia de la retrospectiva del scrum, que en esta nueva etapa de la investigación se propone validar.

PALARAS CLAVE

Desarrollo Software, Metodología, Agiles, Trama vincular, Evaluación.

2.CONTEXTO

El presente trabajo de investigación, en su etapa inicial, se realizó en marco del desarrollo de la tesis doctoral: Modelo de Trabajo Conceptual Referencial para Ingeniería de Procesos. El caso de los Equipos de Desarrollo de Software. Defendida en el Doctorado en Ingenieria, mención Ingenieria Industrial, de la Faciultad de Ingeniería de la Universidad Nacional de Lomas de Zamora. Actualmente se promueve una línea de investigación de continuidad cuyo objeto es realizar una prueba de la herramienta en un conjunto de empresas de un cluster.

1. INTRODUCCION

Las metodologías de desarrollo de software han definido lineamientos respecto de cómo los equipos de trabajo en las organizaciones conforman sus procesos de desarrollo (Rumbaugh, 1996). Las mismas, consisten en una filosofía o paradigma de desarrollo que las sustenta y un conjunto de herramientas, técnicas y métodos que asisten al proceso de desarrollo de software.

Sin embargo, aún se observan resistencias en la adopción de estas metodologías por parte de la industria en forma completa. Esta resistencia está sustentada en la suposición de que se incrementan los costos, o bien se subestiman sus beneficios; suponiendo o malinterpretando su aplicación, utilizando en su reemplazo metodologías más conocidas pero perimidas.

En el año 2001 se formalizó el surgimiento de un conjunto de metodologías basadas en procesos de control empíricos; que sus autores ¨Kent Beck; James Grenning; Robert C. Martin; Alistair Cockburn; Jeff Sutherland; Martin Fowler y otros acordaron en denominar "ágiles". Las metodologías ágiles se definen como un conjunto de lineamientos y mejores prácticas para el desarrollo de software que valoran tanto la propia experiencia como la de otros involucrados en el proceso de desarrollo.

El movimiento ágil no es anti-metodología, de hecho lo que busca es restaurar la credibilidad en la palabra metodología, considerándola como una herramienta que ayuda al proceso y no lo detiene (Agile Alliance, 2001).

Varios de los aspectos que aún las metodologías ágiles no terminan de resolver están vinculados a las relaciones interpersonales entre los diferentes stakeholders. Estos se definen como los grupos interesados y deben ser considerados un elemento esencial en la planificación estratégica

de los negocios (Freeman, 1984); en la trama vincular (Pichon-Rivière, 2002), y en la comunicación.

3. LINEAS DE INVESTIGACION

Teniendo en cuenta las debilidades aún presentes en las metodologías ágiles se han formulado los siguientes interrogantes: ¿Es posible conformar un marco de trabajo que permita mejorar las relaciones interpersonales entre los afectados en un proyecto de desarrollo de software?

¿Son las metodologías ágiles válidas para lograr los objetivos que se propone un equipo de trabajo durante el desarrollo de un producto software?

¿Los recursos con los que cuentan actualmente los equipos de trabajo son suficientes para tener relaciones interpersonales acordes a las exigencias a las que se ven sometidos?

¿Mejorar la comunicación en los equipos de trabajo, sería un aporte a los resultados esperados?

En virtud de este último planteo, las siguientes hipótesis, funcionaron como anticipaciones de sentido para el abordaje de la investigación.

La utilización de metodologías ágiles en el proceso de construcción de software mejora la calidad del producto resultante.

Un software que responde a las necesidades de los usuarios satisfactoriamente (un software con calidad), está relacionado con las competencias de los stakeholders, especialmente de los desarrolladores.

La comunicación entre los stakeholders es un factor que interviene de forma crítica en el desarrollo del software.

La trama vincular en el equipo de desarrollo, incluyendo al dueño del producto (Producto Owner del Scrum) y a los usuarios directos es un factor determinante de la calidad del producto a desarrollar.

Objetivos de la investigación: Indagar cuáles son los métodos de desarrollo utilizados en el mercado, en las localidades maduras de la República Argentina.

Identificar las competencias clave valoradas por los desarrolladores de software de los centros de desarrollo a investigar.

Analizar el papel que desempeña la comunicación entre los stakeholders y la dinámica de los vínculos que se establecen en el equipo de desarrollo.

Diseñar un modelo que se constituya en un marco orientado a definir un conjunto de buenas prácticas conducentes a mejorar el ecosistema de los equipos de desarrollo de software en función del análisis que surge de la investigación.

Validar el modelo mediante su aplicación en diversos casos.

Se destaca como beneficiarios directos los equipos de trabajo que se conforman cada vez que inicia un proyecto para desarrollar un producto de software.

La industria, por otra parte busca permanentemente nuevas metodologías para mejorar su productividad y calidad, en caso de incorporar una nueva metodología cuya adaptación resulte un incremento fácilmente incorporable.

Los lazos que se conforman entre la industria, cada una de sus organizaciones y los equipos de trabajo, eslabón final en esta cadena de valor; tienen visibles necesidades de ser mejorados especialmente en el campo de las relaciones comunicacionales.

4.RESULTADOS ALCANZADOS:

La encuesta fue administrada 60 profesionales y se focalizó en: **competencias técnicas**; Programación o Desarrollo, Manejo de Técnicas de Ingeniería de Requerimientos (Casos de Uso, Técnicas de Elicitación), Gestión ágil de requerimientos (user stories), Testing de Software, Gestión de Personas, Gestión Ágil de Proyectos, Gestión Tradicional de Proyectos, Modelado de Arquitecturas de Software

Diseño de Interacción Humano-Máquina Manejo de Herramientas para Testing Automatizado

Manejo de Herramientas para Gestión de Configuración/Integración Continua

competencias conductuales; Resolución de problemas complejos, Trabajo en Equipo, Autoorganización, Manejo de Inglés, Pensamiento Crítico, Orientación a Servicios, Negociación. Comunicación, Involucramiento, Compromiso, Cooperación

A partir de los resultados obtenidos se propone un modelo de trabajo para ingeniería de procesos y en particular para todos los proyectos de desarrollo de software, con el objetivo de ayudar

a los profesionales a evaluar la salud del grupo como equipo de trabajo, comprometido en la creación de software y en mejorar su forma de vincularse entre todos los involucrados directos e indirectos. El modelo busca reforzar aspectos culturales que conlleven a mejorar las estructuras organizacionales, dado que a pesar del hecho que, según Larman (Larman, 2017) plantea en uno de sus postulados conocidos como las leyes de Larman, "la cultura sigue a la estructura", es decir el mecanismo de funcionamiento del equipo es un reflejo del mecanismo de funcionamiento de la organización. La propuesta apunta a, en un sentido inverso, trabajar sobre los equipos para provocar un cambio y una mejora que se traslade a las organizaciones (botton up). Analizando el comportamiento de los grupos que adhirieron a prácticas ágiles, en su mayoría se observan aspectos "débiles" en su adopción. Debido a aspectos de la trama vincular, ya que la misma aparece como superficial, subestimada, de forma tal que se toman decisiones "no-ágiles". Es preciso considerar que en estas metodologías la producción depende del trabajo en equipo y no de las individualidades; por lo que banalizar la trama vincular conspira contra la productividad.

El modelo de trabajo se manifiesta en la generación de heurísticas y de indicadores clave de vínculos. Las heurísticas está agrupadas en ejes temáticos los cuales se enumeran a continuación: comunicación entre los stakeholders, clima laboral, tele entre los miembros del grupo, pertenencia, pertinencia y cooperación.

Asímismo, los indicadores clave de vínculos para un equipo de proyecto de desarrollo de software, respetan los mismos ejes temáticos de las heurísticas y constituyen la herramienta de evaluación de su grado de presencia en los equipos de trabajo. A continuación se presentan las heurísticas, organizadas por ejes temáticos enfocadas a los vínculos que se establecen en los equipos de trabajo. Se generó una rúbrica que permita evaluar el índice en cada eje temático. La rúbrica es un instrumento que sirve para evaluar el eje en función de los indicadores de vínculos, teniendo en cuenta para ello un conjunto de criterios que han sido acordados en forma previa a su utilización por parte del equipo. En la rúbrica, se define el valor alcanzado por el grupo en cada eje de acuerdo a un indicador clave de vinculo, una escala nominal: nulo impedido, transición, sostenible e ideal. Esto permite evaluar al grupo de trabajo en el proceso y al finalizar un ciclo completo. La propuesta recomienda la inclusión de las prácticas en forma permanente e integrada al trabajo de los equipos de desarrollo. Los indicadores clave de vínculo utilizados a partir de un framework ágil para la gestión de proyectos, en el contexto de la ceremonia que el equipo se reserva para inspeccionar y adaptar sus propias prácticas, conocida comúnmente como Retrospectiva, la cual parte de parte de Scrum.

Esta ceremonia es la materialización del último principio del manifiesto ágil que plantea que: "A intervalos regulares el equipo reflexiona sobre cómo ser más efectivo para a continuación ajustar y perfeccionar su comportamiento en consecuencia".

Tal como define Scrum, la retrospectiva es la última ceremonia que se realiza antes de cerrar el Sprint y brinda la oportunidad de reforzar y afianzar buenas prácticas y mejorar aquellas que no lo son tanto.

En este sentido y sustentado en la información obtenida del procesamiento de las encuestas, se observó que el foco de la Retrospectiva en la mayoría de los equipos está en mejorar aspectos relacionados con las competencias técnicas, desplazando las competencias conductuales que son el núcleo del modelo conceptual referencial propuesto.

El modelo que se presenta propone fortalecer la ceremonia del Scrum denominada *retrospectiva*, con la inclusión del análisis de los ejes temáticos y las heurísticas propuestas; incorporando en ese momento la reflexión y medición de la trama vincular del grupo operativo. Se propone la incorporación de una dinámica, en la cual, teniendo la tabla de referencia en forma visible y el framework de trabajo sobre el cual el grupo irá marcando el estado logrado al final de cada sprint.

Al sprint siguiente el grupo realizará una vez más, en la misma ceremonia (retrospectiva), la misma dinámica. Al finalizar el proyecto, se propone una ceremonia final denominada

"Retrospectiva Recargada" en la cual el grupo en un proceso de discernimiento evalúe la evolución en base al conjunto de frameworks y diagramas y pueda tomar acciones de mejora para futuros proyectos.

Indicador clave de vinculo → nulo impedido, transición, sostenible e ideal.	
Eje temático ↓	**Heurística**
Comunicación entre los stakeholders	Visión compartida
	Reuniones como herramienta de creación de valor
Clima de trabajo	Gestión del tiempo/Foco en la tarea/Uso de espacios compartidos
Tele entre los miembros del grupo	Respeto por el grupo
	Capacidad de Expresión
	Construcción de Equipo
Pertenencia	Objetivos compartidos
	Compromiso
	Conducta activa
Pertinencia	Identificación con el equipo
	Equipo Involucrado
	Equipo Integrado
Cooperación	Colaboración
	Asistencia /Acompañamiento
	Solicitud de asistencia

5. RESULTADOS ESPERADOS

Se espera validar la Rúbrica de Evaluación del funcionamiento de los equipos de desarrollo software, a partir de su implementación en un conjunto de casos.

6. FORMACION 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, dos tesis de Maestría, así como para tesinas de grado.

Bibliografía

Yourdon, E. (1982). *Análisis Estructurado Moderno.* Prentice Hall Ispanoamericana S.A.

Rumbaugh, J. (1996). *Modelado y Diseño Orientado a Objetos.* Prentice Hall.

Jacobson, I. (1994). *Object Oriented Software Engineering.* Estados Unidos: Addison Wesley.

Booch, G. (1991). *Object OrientedDesign with Applications.* California: Benajmin / Cummings.

Agile Alliance. (2001). *The agile manifesto*. Obtenido de http://www.agilealliance.org/the-alliance/the-agile-manifesto/

Fowler, M. (13 de 12 de 2005). *http://martinfowler.com/articles/newMethodology.html.* Obtenido de http://dcc.uab.es/sites/default/files/21290/teoria/Fow2005-TheNewMethodology.pdf

Freeman, E. (1984). *Strategic Management: A Stakeholdrs Approach.* Boston: Pitmaan Press.

Pichon-Rivière, E. (2002). *Teoría del Vínculo* (1ra. reimpresión 24 ed.). Buenos Aires: Nueva Visión.

Larman, C. (07 de 03 de 2017). *Larman's Laws of Organizational Behavior*. Obtenido de http://www.craiglarman.com/wiki/index.php?title=Larman%27s_Laws_of_Organizational_Behavior

MAXWELL, J. A. (1996). *QUALITATIVE RESEARCH DESIGN. An Interactive Approach.* California: Sage Publicatios.

Sampieri, R. H. (2014). *Metodología de la Investigación* (6 edición ed.). Mexico D.F.: Mc Graw Hill Education.

Kuechler, M. (1988). *The Survey Method: An Indispensable Tool for Social Sciences Everywhere?* Florida: American Behavioralist Scientist.

Riviere, E. P. (1997). *El Proceso Grupal: del Psicoanalisis a la Psicologia Social.* Buenos Aires: Coleccion Psicologia Contemporanea.

Lema, V. Z. (1986). *Conversaciones con Enrique Pichon-Riviere.* Buenos Aires: Ediciones Cinco S.A.

Servicios Web que integran aplicaciones de aprendizaje ubicuo

Coali, Nicolás; Únzaga Silvina; Alvarez, Margarita; González, Gabriela; Fernández Reuter, Beatriz; Duran, Elena

Instituto de Investigaciones en Informática y Sistemas de Información
Departamento de Informática - Facultad de Ciencias Exactas y Tecnologías
Universidad Nacional de Santiago del Estero (UNSE)
e-mail: nico.coali@gmail.com; sunzaga2016@gmail.com; alvarez@unse.edu.ar; ggonzalez@unse.edu.ar; bfreuter@unse.edu.ar; eduran@unse.edu.ar

RESUMEN

La emergencia de nuevos paradigmas en la Web con sus consiguientes aplicaciones en el campo educativo y, el gran desarrollo de la tecnología en comunicaciones y de los dispositivos móviles inteligentes, han dado lugar al surgimiento del aprendizaje ubicuo.

Con este aprendizaje se superan las limitaciones de ambiente de aprendizaje tradicional, y lo extienden haciendo realidad la idea de aprendizaje en cualquier lugar y en cualquier momento, permitiendo a las personas acceder a mejores experiencias de aprendizaje.

En el marco del proyecto de investigación "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas", se han propuesto cinco servicios que se ofrecen a los estudiantes en un ambiente de aprendizaje ubicuo, de los cuales dos de ellos se han desarrollado independientemente.

Para solucionar esta dificultad y debido a la necesidad de contar con una única aplicación, en esta línea de investigación se realizará la integración empleando servicios Web.

***Palabras clave*:** Servicios Web, Integración, Aplicaciones de aprendizaje ubicuo.

CONTEXTO

Este trabajo se desarrollará en el marco de una Beca del Programa de Beca CICyT –UNSE, Disposición Nº: 133/18. Además, se inserta en el proyecto "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas", correspondiente a la convocatoria 2017-2020 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SICYT - UNSE).

1. INTRODUCCIÓN

En los últimos años, gracias a los avances de los dispositivos móviles, sensores y las telecomunicaciones ha surgido una nueva área: computación ubicua, que en el campo de la educación ha dado origen al aprendizaje ubicuo.

Este aprendizaje más que un fenómeno de la tecnología es una visión que desafía a revisar el entorno educativo para proponer nuevos escenarios educativos, ya que permiten poner a disposición de los estudiantes una gran diversidad de recursos educativos, crear nuevos y variados entornos de formación, personalizar el aprendizaje y posibilitar la realización de un conjunto de actividades formativas desde cualquier momento y lugar y, desde cualquier dispositivo (Saadiah, 2010).

El desarrollo de aplicaciones de aprendizaje ubicuo es una tarea compleja, debido a que se requiere considerar las características y necesidades de los usuarios, múltiples formas de movilidad, la tecnología móvil, la diversidad de los transportistas, y también, la diversidad de escenarios de aprendizaje que pueden presentarse.

Debido a esta complejidad, en el marco del proyecto de investigación "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas" se ha propuesto una arquitectura de una aplicación ubicua (Durán et al., 2014), que ofrece diferentes servicios de aprendizaje ubicuo:

- *Servicio 1:* Recomendar Objetos de Aprendizaje para puntos de interés seleccionados por el estudiante (Loto y Durán, 2015), (Durán et al., 2016).

- *Servicio 2*: Recomendar un camino de aprendizaje a partir de un objetivo de aprendizaje seleccionado por el estudiante.
- *Servicio 3*: Recomendar puntos de interés cercanos a la ubicación del estudiante (Montoto et al., 2017).
- *Servicio 4*: Recomendar expertos para que asesoren al estudiante para cumplir con una tarea (Durán y Alvarez, 2017), (Fernandez Reuter et al., 2018).
- *Servicio 5*: Recomendar pares para que asesoren al estudiante para cumplir con una tarea (Alvarez et al., 2017).

A la fecha dos de estos servicios (1 y 3) fueron desarrollados como aplicaciones de aprendizaje ubicuo independientes. Dada la necesidad de integrarlos, en esta línea de investigación se propone utilizar servicios Web para este fin.

Los servicios Web son servicios autónomos e independientes que se ofrecen mediante la Web. Principalmente, permiten que las aplicaciones sean más modulares y desacopladas, facilitando su reutilización en distintas plataformas o lenguajes de programación.

Estos servicios aportan interoperabilidad entre aplicaciones de software independientemente de sus propiedades o de las plataformas sobre las que se instalen.

Los servicios Web fomentan los estándares y protocolos basados en texto, que hacen más fácil acceder a su contenido y entender su funcionamiento. Permiten que servicios y software de diferentes compañías ubicadas en diferentes lugares geográficos puedan ser combinados fácilmente para proveer servicios integrados.

Entre las principales razones para el uso de servicios Web es que se pueden utilizar con HTTP sobre Transmission Control Protocol (TCP) en el puerto de red 80. Dado que las organizaciones protegen sus redes mediante firewalls (que filtran y bloquean gran parte del tráfico de Internet), cierran casi todos los puertos TCP salvo el 80, que es, precisamente, el que usan los navegadores Web. Los servicios Web utilizan este puerto, por la simple razón de que no resultan bloqueados. Es importante señalar que los servicios Web se pueden utilizar sobre cualquier protocolo, sin embargo, TCP es el más común.

Otra razón es que, antes de que existiera SOAP (Simple Object Access Protocol), no había buenas interfaces para acceder a las funcionalidades de otras computadoras en red. Las que había eran ad hoc y poco conocidas, tales como Electronic Data Interchange (EDI), Remote Procedure Call (RPC), u otras API.

Una tercera razón por la que los servicios Web son muy prácticos es que pueden aportar gran independencia entre la aplicación que usa el servicio Web y el propio servicio. De esta forma, los cambios a lo largo del tiempo en uno no deben afectar al otro. Esta flexibilidad será cada vez más importante, dado que la tendencia a construir grandes aplicaciones a partir de componentes distribuidos más pequeños es cada día más utilizada (Miranda Pons et al., 2014).

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Con esta línea de investigación se busca realizar contribuciones conceptuales y metodológicas en el campo de los servicios Web. En consecuencia, el objetivo general es:

Integrar las aplicaciones de aprendizaje ubicuo desarrolladas en el marco del proyecto "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas" utilizando servicios Web.

Además, se han definido los siguientes objetivos específicos:

- Analizar y evaluar los distintos métodos.
- Analizar y evaluar los diferentes estándares para el diseño de Servicios Web.
- Definir la arquitectura orientada a servicios que permita la integración de las aplicaciones.
- Diseñar, construir y evaluar los servicios Web correspondientes.
- Integrar las aplicaciones ya desarrolladas en el marco del proyecto, correspondientes a los servicios 1 y 3 utilizando los servicios Web construidos.

Con el propósito de dar cumplimiento a los objetivos planteados, se realizarán las siguientes actividades:

1. *Exploración e investigación bibliográfica*: a partir del estudio exploratorio se desarrollarán los marcos conceptual y metodológico que permitirán abordar el diseño y la construcción del prototipo de integración de servicios Web. Se seleccionará la metodología, arquitectura, técnicas y herramientas para el desarrollo del mismo. Se utilizarán, principalmente, las técnicas de estudio de documentación, revisión y fichaje bibliográfico.
2. *Diseño de la arquitectura orientada a servicios*: se analizarán las diferentes arquitecturas para integrar los servicios Web a desarrollar. Se diseñará la arquitectura de integración de servicios.
3. *Diseño de los servicios Web*: se realizará el diseño de los servicios Web a construir. Para ello se utilizarán técnicas provenientes de la Ingeniería de Software y de la Ingeniería Web.
4. *Construcción de los servicios Web*: a partir de los modelos de diseño, obtenidos de la actividad anterior, se procederá a desarrollar, en los lenguajes de programación pertinentes. Se realizarán las pruebas de unidad y las pruebas de integración para corroborar la correcta construcción de la aplicación. En base a los resultados de la evaluación se realizarán ajustes en los módulos y/o en la integración de los mismos, si fuera necesario.
5. *Ejecución de la integración de las aplicaciones:* se implementará el prototipo probando la integración de las aplicaciones ya desarrolladas en el marco del proyecto.
6. *Análisis de resultados*: se procesarán los datos recolectados calculando las métricas previamente definidas. Se elaborará una representación tabular y gráfica de los resultados obtenidos, acompañando los mismos con un análisis interpretativo de los resultados.

4. RESULTADOS ESPERADOS

Los resultados que se esperan obtener de esta investigación, conforme a los objetivos, son:

- Estado del arte referido a los diferentes estándares para el diseño de Servicios Web.
- Arquitectura orientada a servicios de la aplicación ubicua integral.
- Servicios Web correspondientes a las aplicaciones ubicuas individuales desarrolladas en el marco del proyecto.
- Prototipo que integre los servicios.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea se integra por un Becario del Programa de Beca CICyT – UNSE, un Director de la Beca y Director del Proyecto de Investigación indicado en el contexto.

En esta línea de investigación, el becario desarrollará su trabajo final para optar por el título de Licenciado en Sistemas de Información de la Universidad Nacional de Santiago del Estero.

5. BIBLIOGRAFÍA

Alvarez M., Únzaga S. y Durán E. (2017). Método para generar recomendaciones personalizadas para integrar grupos de aprendizaje ubicuo y colaborativo. XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). Libro de actas. Pag. 363- 372. ISBN 978-950-34-1539-9. Compiladores: De Giusti, Armando Eduardo y Pesado, Patricia Mabel. Disponible en: http://sedici.unlp.edu.ar/handle/10915/63019 La Plata. Prov. Buenos Aires. 9 al 13 de octubre.

Durán, E.; Álvarez, M. y Únzaga, S. (2014). Ontological model-driven architecture for ubiquitous learning applications. EATIS 2014 - 7th Euro American Association on Telematics and Information Systems. Valparaíso (Chile). 2 al 4 de Abril de 2014.

Durán, E., Únzaga S. y Álvarez, M. (2016). Instanciación del Modelo de Servicios para una aplicación de apoyo al Aprendizaje Ubicuo en un curso de Redes de Computadoras. Revista TEKNOS de la Facultad de Ingeniería de la Fundación Universitaria Tecnológico Comfenalco. ISSN 1900–7388.

Durán E. y Álvarez M. (2017). Method for Generating Expert Recommendations to Advise students on Ubiquitous Learning Experiences. Conferencia Internacional de la Sociedad Chilena de Ciencia de la Computación (SCCC 2017). Arica, Chile. 16 al 20 de Octubre de 2017.

Fernández Reuter B., Álvarez M., Durán E. y González G. (2018). Multi-agent system model for tutor recommendation in ubiquitous learning environments. Proceedings First Workshop on Advanced Virtual Environments and Education - WAVE2 2018. Pag. 10-17. Florianópolis, Brasil, del 04 al 05 de Octubre de 2018. DOI: 10.5753/wave.2018.1. Disponible en https://mafiadoc.com/wave2-2018_5c1e5c4f097c4776518b464c.html.

Loto M. y Durán E. (2015). Diseño de una aplicación móvil personalizada de apoyo al aprendizaje de Redes de Computadoras. X Congreso de Tecnología en Educación & Educación en Tecnología. Corrientes.

Miranda Pons, Y., Santiesteban, I., Gari, M., Rodríguez Soberats, D. (2014) Propuesta de integración de servicios web para las aplicaciones del centro FORTES. https://www.researchgate.net/publication/275580771

Montoto S., Únzaga S. y Durán E. (2017). Diseño de una aplicación ubicua personalizada de apoyo a estudiantes del curso de ingreso. Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. San Fernando del Valle de Catamarca. 10 y 11 de Agosto. ISSN 1853-6662. pp: 191-198.

Saadiah Y., Erny Arniza A.and Kamarularifin A. J. (2010). The definition and characteristics of ubiquitous learning: A discussion. In Proceedings of International Journal of Education and Development using Information and Communication Technology (IJEDICT), Vol. 6, Issue 1, 117-127.

Sistemas móviles multiplataforma con realidad aumentada. Aplicaciones en educación y salud.

Susana I. Herrera, Marta C. Fennema, María I. Morales, Marilena Maldonado, Rosa A. Palavecino, Federico Rosenzvaig, Adriana Macedo, Claudia Suárez, Ricardo Pellicer, Rita Villavicencio

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
{sherrera, imorales, rosypgg, marilena, frosenzvaig}@unse.edu.ar, crisfen@yahoo.com, {adrumacedo, ricardopellicer}@gmail.com, {cfemalekeypal, crvillavicencio}@hotmail.com

Resumen
En este trabajo se presenta una investigación que constituye una continuación de los proyectos sobre Computación Móvil, llevados a cabo en la Universidad Nacional de Santiago del Estero (UNSE) en 2012-2016 y 2017-2018 en los cuales se lograron resultados vinculados a la eficiencia de aplicaciones móviles, el aprendizaje basado en dispositivos móviles (m-learning), realidad aumentada (RA), redes móviles, herramientas de visualización y método para el desarrollo móvil multiplataforma con RA.
Sobre la base de dichos resultados, se propone investigar métodos y herramientas para el desarrollo/mantenimiento de aplicaciones móviles multiplataforma, abarcando: a) aplicaciones con RA generadas con herramientas de compilación cruzada y b) aplicaciones web responsivas. Se construirán prototipo de aplicaciones para aprendizaje de niños hipoacúsicos con implantes cocleares (basados en prácticas de entrenamiento auditivo) y así como también para estudiantes de escuelas ordinarias.

***Palabras clave*:** Aplicaciones móviles multiplataforma, realidad aumentada, m-learning, aplicaciones móviles para niños hipoacúsicos.

1 Contexto

Esta investigación se lleva a cabo en el marco del proyecto denominado "*Sistemas móviles multiplataforma con realidad aumentada. Aplicaciones en educación y salud*", del Instituto de Investigaciones en Informática y Sistemas de Información (IIISI), financiado por el Consejo de Ciencia y Técnica de la UNSE, durante el período enero 2019-diciembre 2020. Es una continuación de los proyectos "*Optimización de la calidad de los Sistemas Móviles mediante la implementación de nuevas arquitecturas, Realidad Aumentada, técnicas de visualización y redes móviles Ad-Hoc. Aplicaciones en m-learning y en gestión del conocimiento*", implementado entre los años *2012* y 2016 y "*Computación móvil: desarrollo de aplicaciones móviles y análisis forense"* entre los años 2017 y 2018.
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.
En cuanto al dominio de aplicación, se consideran las áreas Educación y Salud. En relación a Educación, se diseñarán prácticas de m-learning para nivel secundario y universitario. En el nivel universitario, se aplica en las unidades académicas Facultad de Ciencias Exactas y Tecnologías (FCEyT) y la Facultad de Humanidades, Ciencias Sociales y de la Salud, ambas de la UNSE. En el nivel secundario, se trabaja en la Escuela Técnica Nº 6 de La Banda y en escuelas secundarias de Cali, Colombia (en el marco del convenio de colaboración firmado entre la FCEyT y la Universidad Unicatólica de Cali).
En relación al área Salud se trabajará con niños hipoacúsicos, trabajando en el marco de un convenio específico de colaboración firmado entre la FCEyT y el Servicio de Detección Temprana de Hipoacusia del Ministerio de Salud de la Provincia de Santiago del Estero.

2 Introducción

El tema principal de esta investigación es la Computación Móvil, la cual se apoya en tres ejes elementales (Talukder et al., 2010): dispositivos móviles, redes móviles y aplicaciones móviles. Esta investigación abarca el eje referido a aplicaciones móviles, abordadas desde la Ingeniería del Software.

A continuación, se mencionan los principales resultados alcanzados por este equipo de investigación en los proyectos anteriores de Computación Móvil. Dado que es lo que sustenta este nuevo proyecto que se desarrollará en 2019-2020.

Se estudiaron las arquitecturas de aplicaciones móviles propuesta por Challiol (2012): cliente, híbrida o web. Najar (2013) ha diseñado modelos de arquitecturas móviles, tanto web como híbridas, orientadas a servicios, para SO Android.

Con posterioridad, el aumento del uso de las aplicaciones móviles ejecutadas bajo diversos SO móviles (Android, iOS, Windows), hizo que el paradigma de programación de aplicaciones cambie hacia el desarrollo multiplataforma: generar aplicaciones móviles que puedan ser ejecutadas por diversos sistemas operativos. En Delía (2017) y Delía et al. (2018), se sostiene que las aplicaciones móviles multiplataforma pueden ser: nativas generadas con proyectos diferentes de acuerdo al SO, webs responsivas, híbridas, interpretadas y generadas por compilación cruzada. El proyecto 2017-2018, siguiendo esta clasificación, ha abordado el desarrollo de aplicaciones móviles multiplataforma, generando aplicaciones web responsivas, híbridas (usando Ionic) y por compilación cruzada (usando Xamarin). Estas últimas involucró el desarrollo de interfaces con RA basada en objetos 3D animados.

Respecto a los métodos de desarrollo de software, se estudiaron y/o utilizaron métodos ágiles como eXtreme Programing (XP), Agile Unified Process (AUP), Mobile-D (Abrahamsson et al., 2004). En base a Mobile-D, en el proyecto 2017-2018 se desarrolló una propuesta metodológica denominada Mobile-RA, específica para el desarrollo móvil con RA (Budán, 2018).

En relación a la RA, cabe mencionar que 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 computadora o algún dispositivo móvil (Azuma, 2001). Los contenidos que muestran las aplicaciones de RA pueden ser: texto, imágenes, audio, video y objetos 3D. Krange et al. (2002) recalcan la importancia de los objetos 3D en el aprendizaje, ya que proveen: sensación de tocar, sentido, orientación y posición en el espacio. La RA es atractiva, dado que se alinea con el aprendizaje activo (Moralejo, 2014). Existen dos tipos de RA: geolocalizada y basada en marcadores. En el grupo de investigación se trabaja con Nivel 0 de RA (Prendes Espinosa, 2015): visualización de objetos 3D a partir de la lectura de marcadores. Cuando se reconoce un patrón en particular, en su posición, se superpone una imagen digital en la pantalla.

El m-learning es el proceso de adquirir conocimiento mediante una relación dialógica entre el entorno y las personas y/o las personas entre sí, a través de una mediación con tecnología móvil, tanto en contextos de aprendizaje formales como no formales e informales; involucra en el aprendiz competencias tecnológicas para manipular los dispositivos móviles, competencias relacionadas con el aprendizaje autónomo y con la capacidad de interacción y comunicación (Herrera, 2016).

Las características principales del m-learning son: el ecosistema (Woodill, 2011), los modos de interacción (Woodill, 2011), los enfoques para la implementación de prácticas (Pachler et al., 2012), la relación con la vida cotidiana y cuestiones generales de aprendizaje que son importantes para el diseño de prácticas de m-learning.

Este grupo de investigación ha desarrollado aplicaciones móviles para aprendizaje y se diseñaron e implementaron diversas prácticas de m-learning en instituciones de diversos niveles educativos, abarcando el aprendizaje en Tecnologías, Programación, Matemática, Inglés y Sistemas (Maldonado et al., 2018; Morales et al.,2018; Herrera et al., 2018;

Herrera et al., 2017; Herrera et al., 2014, Palavecino et al.,2016). Las prácticas se diseñaron usando un marco sistémico y ecológico para el análisis, diseño y evaluación de experiencias de m-learning, denominado MADE-mlearn (Herrera et al., 2013).

Teniendo en cuenta todos los antecedentes mencionados, en este proyecto se desarrollarán aplicaciones para niños hipoacúsicos. Varios autores en el campo de la informática se han dedicado a abordar este tema; entre ellos se destacan Pérez-Arévalo et al. (2017), Manresa-Yee et al. (2017), y Peñeñory et al. (2010). La hipoacusia es la discapacidad auditiva. La cóclea (o caracol) se encuentra en el oído interno y deficiencias en ella no sólo produce pérdida auditiva sino también dificultades para mantener el equilibrio físico de la persona.

A partir del uso de implantes cocleares en niños hipoacúsicos, se intenta que los niños implantados desarrollen el lenguaje a temprana edad y se inserten en la educación estándar. Para lograr esto, se requiere que el niño tenga una rehabilitación robusta que involucre a la familia y a profesionales como fonoaudiólogos, audiólogos, logopedas. En este contexto, el uso de dispositivos móviles para la rehabilitación de estos niños es crucial, dado que se trata de dispositivos de la vida cotidiana de bajo costo.

Existen aplicaciones móviles provistas por algunas compañías de fabricación de implantes cocleares (Ej. Coclenado App). Estas contienen ejercicios de identificación, discriminación y comprensión; pero no están orientadas al aprendizaje de conceptos que promuevan la inserción del niño en el nivel educativo preescolar estándar. Además, las interfaces visuales no son dinámicas, resultando poco atractivas.

A partir de la situación presentada, se considera relevante diseñar aplicaciones móviles multiplataforma con RA para el aprendizaje de niños con implantes cocleares.

3 Líneas de investigación y desarrollo

Se propone investigar sobre métodos y herramientas para la generación de aplicaciones móviles multiplataforma, abarcando: a) aplicaciones web responsivas, b) aplicaciones con RA híbridas, interpretadas y generadas por compilación cruzada.

Se ampliarán las aplicaciones educativas desarrolladas en las investigaciones anteriores y se desarrollará una aplicación para niños hipoacúsicos.

Se propone también avanzar en la investigación en m-learning, respecto al cual se pretende diseñar nuevas prácticas colaborativas y/o basadas en RA usando MADE-mlearn, desarrollar un repositorio de Objetos 3D para uso en aplicaciones móviles educativas con RA y, también, diseñar Objetos de Aprendizaje para contextos móviles.

Considerando la amplitud y relevancia de aspectos de Computación Móvil que se abordarán, esta investigación se organiza en las siguientes líneas:

- Desarrollo de aplicaciones móviles multiplataforma con RA.
- Optimización de aplicaciones web responsivas.
- M-learning.

4 Resultados esperados

Se espera generar nuevo conocimiento científico-tecnológico en el área de la Computación Móvil, relativos a:

- Ingeniería del Software: métodos y herramientas para desarrollar aplicaciones móviles multiplataforma con y sin RA.
- Productos software: se ampliará la aplicación móvil con RA de AlgeRA, Se internacionalizará la aplicación web responsiva ImaColab y se desarrollará una nueva aplicación móvil multiplataforma para Rehabilitación de Niños Hipoacúsicos.
- M-learning: se diseñarán e implementarán experiencias, con y sin RA, y sus impactos en la motivación intrínseca. Se optimizarán las aplicaciones de m-learning AlgeRA e ImaColab.

5 Formación de recursos humanos

La concreción de este proyecto contribuirá a la formación y capacitación de los investigadores, estudiantes de grado y de posgrado involucrados. Incentivará a los alumnos a iniciarse en las actividades de investigación y

favorecerá la realización de sus trabajos finales de grado y posgrado en las líneas de investigación de este proyecto.

Esta investigación contribuye al desarrollo de una tesis doctoral (UNLP-Ecuador), una tesis de Maestría en Informática Educativa (UNSE), una tesis de Especialización en Enseñanza en Tecnologías (UNSE) y dos trabajos finales de grado de la Licenciatura en Sistemas de Información (UNSE).

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. Abrahamsson, P., Hanhineva, A., Hulkko, H., Ihme, T., Jäälinoja, J., Korkala, M.& Salo, O. (2004). Mobile-D: an agile approach for mobile application development. In Companion to the 19th annual ACM SIGPLAN conference on Object-oriented programming systems, languages, and applications (pp. 174-175). ACM.
2. Azuma, R. (2001). Augmented Reality: Approaches and Technical Challenges. Fundamentals of Wearable Computers and Augmented Reality. Mahwah, New Jersey: W. Barfield, Th. Caudell.
3. Budán, P., Rosenzvaig, F., Najar P., Saavedra, E, Herrera, S., Morales, M. Mobile-RA: Método Ágil para el Desarrollo de Aplicaciones Móviles Multiplataforma. CONAIISI 2018. Actas del Congreso. ISSN 2347-0372. Noviembre 2018. Mar del Plata, Argentina.
4. Challiol, C. (2012). Apuntes de Curso de Posgrado sobre Computación Móvil. Universidad Nacional de Santiago del Estero. Secretaría de Extensión de la Facultad de Ciencias Exactas y Tecnologías.
5. Delía 2017
6. Delía 2018
7. Herrera, S., Sanz, C., Morales, M., Palavecino, R., Maldonado, M., Irurzun, I. Carranza, A., Suárez, G. (2018). M-learning con Realidad Aumentada basada en Objetos 3D. Workshop de Investigadores en Ciencias de la Computación, WICC. ISBN 978-987-3619-27-4. Corrientes.
8. Herrera, S., Morales, M., Palavecino, R., Maldonado, M., Irurzun, I. Carranza, J. (Herrera et al., 2017). M-Learning con Realidad Aumentada. I Jornada de Vinculación con el medio, difusión de resultados de investigación y desarrollos en informática. Santiago del Estero, Argentina, 2017.
9. Herrera, S. (2016). MADE-mlearn: Marco para el Análisis, Diseño y Evaluación de experiencias de m-learning. Tesis Doctoral. Universidad Nacional de La Plata. La Plata, Argentina.
10. Herrera, S., Morales, M., Sanz, C., Fennema, C. (2014). Aprendizaje Basado En Dispositivos Móviles. Santiago del Estero: EDUNSE - Editorial de la UNSE. pag.196. isbn 978-987-1676-18-7.
11. 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.
12. Krange, I., Fjuk, A., Larsen, A., Ludvigsen, S. (2002). Describing construction of. In Proceedings of the Conference on Computer (págs. 82-91). CSCL Community.
13. Maldonado, M., Herrera, S., Irurzun, I., Carranza, J., Palavecino, R., Macedo, A., Suárez, C. (2018). M-learning: Aprendizaje en Diversos Niveles Educativos Usando ImaColab. Investigaciones en Facultades de Ingeniería del NOA. Libro de Resúmenes. ISSN 2618-396X. Ed. UNSE. Santiago del Estero, septiembre 2018.
14. Manresa-Yee, C., Mas-Sansó, R. and Cano, S. (2017). Juego serio para entrenar la discriminación y la identificación auditiva en niños con discapacidad auditiva. In Proceedings of International Conference on Human Computer Interaction, Cancun, Quintana roo Mexico, September 2017 (Interaccion'17). 6 pages.

15. Moralejo, L. (2014). Análisis comparativo de herramientas de autor para la creación de actividades de realidad aumentada. La Plata.
16. Morales, M., Herrera, S., Maldonado, M., Budán, P., Rosenzvaig, F. (2018). M-learning con Realidad Aumentada para el aprendizaje significativo en Álgebra Lineal. Revista Tecnologias na Educacao. Ano10-Número/vol24-Edição SITED 2018. ISSN 1984-4751. Disponible en: http://tecedu.pro.br/ano10-numero-vol24-edicao-tematica-vii/. Brasil.
17. Najar, P. J. (2013). Prototipo de sistema móvil para e-turismo. Trabajo final de Licenciatura en Sistemas de Información. Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero.
18. Pachler, N., Seipold, J., Bachmair, B. (2012). Mobile Learning: Some Theoretical and Practical Considerations. In The "My Mobile" Handbook. Guidelines and Scenarios for Mobile Learning in Adult Education, K. Friedrich, M. Ranieri, N. Pachler, and P. de Theux, Eds. 2012, pp. 11–16.
19. 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.
20. Peñeñory, V., Manresa-Yee, C., Riquelme, I., Collazos, C. and Fardoun, H. (2010). Review of systems to train psychomotor skills in hearing impaired children Conference'10, Copyright 2010 ACM 1-58113-000-0/00/0010. DOI: http://dx.doi.org/10.1145/12345.67890.
21. Pérez-Arévalo, C., Manresa-Yee, C. and Peñeñory, V. (2017). Game to Develop Rhythm and Coordination in Children with Hearing Impairments. In Proceedings of International Conference on Human Computer Interaction, Cancun, Quintana roo Mexico, September 2017 (Interaccion'17), 4 pages. DOI: https://doi.org/10.1145/3123818.3123853.
22. Prendes Espinosa, C. (2015). Augmented Realityand Education: Analysis Of Practical Experiencies. Píxel-Bit. Revista de Medios y Educación. Nº 46. Enero 2015. ISSN: 1133-8482.e-ISSN: 2171-7966. doi: http://dx.doi.org/10.12795/pixelbit.2015.i46.12.
23. Talukder, A., Ahmed, H., & Yavagal, R. (2010). Mobile Computing: Technology, Applications, and Service Creation. Second Edition. -: McGraw-Hill Professional.
24. Woodill, G. (2011). The mobile learning edge. Ed. Mc Graw Hill. ISBN: 978-0-07-173984-9.

TÉCNICAS PARA LA MEJORA DE LA CALIDAD EN LA INGENIERÍA DE REQUISITOS EN LAS EMPRESAS DE SOFTWARE DE ARGENTINA

María Laila Becker, Rafael Blanc, Carlos Casanova, Julián Escalante, Lourdes Pralong, Laura Elena Ríos, Rossana Sosa Zitto
Facultad de Ciencia y Tecnología, Universidad Autónoma de Entre Ríos
rafaellujanblanc@yahoo.com.ar; carlos.casanova16@gmail.com; mdlpralong@gmail.com; elenrios2@gmail.com; rossana.e.sosa.zitto@gmail.com

RESUMEN

La Ingeniería de Requisitos, es una fase crucial en el proceso de desarrollo de software, ya que ésta determina la calidad del diseño y la implementación y tiene un alto impacto en el éxito de los proyectos. Por tal motivo es necesario implementar estrategias que garanticen su éxito, entre ellas técnicas específicas para gestionar el conocimiento, que apunten a mejorar la calidad de los requisitos. Uno de los mayores inconvenientes que se presenta en esta fase es la brecha de interpretación que se produce entre el ingeniero de requisitos y el usuario del software lo que lleva a errores de traducción de necesidades. Para estructurar y categorizar la masa de información proveniente del proceso de educción de requisitos, es posible utilizar técnicas de representación del conocimiento a los efectos de facilitar la comprensión del problema manifestado por el usuario y reducir la brecha conceptual, entre los procesos de educción de requisitos y su modelado.

Palabras clave: Ingeniería de Software, Representación del conocimiento, Ingeniería del Conocimiento, Inteligencia Artificial.

CONTEXTO

El presente trabajo se desarrolla en el ámbito del Grupo de Investigación de Ingeniería de Software (GIISW), perteneciente a la Sede Concepción del Uruguay de la Facultad de Ciencia y Tecnología dependiente de la Universidad Autónoma de Entre Ríos.

1. INTRODUCCIÓN

Actualmente, la industria nacional del sector SSI (Software y Servicios Informáticos), al igual que otras industrias competitivas, es uno de los sectores de la economía que muestra mayor nivel de dinamismo y, por ende, es considerado por el gobierno como un actor de importancia estratégica para el desarrollo del país en su conjunto. [1], [2].

Dentro de esta industria uno de los retos fundamentales de las empresas desarrolladoras de software es implementar aplicaciones que sean entregadas a tiempo, que no involucren presupuestos elevados y que satisfagan las necesidades del usuario, utilizando para su efecto metodologías y herramientas que guíen el proceso de desarrollo de Software [3].

Otra exigencia que ocupa a la industria de software es la calidad, este tema día a día genera mayor interés, así lo demuestra el informe presentado en 2017 por la OPSSI (Observatorio permanente de la industria del software y servicios informáticos). Uno de los principales motivos es que las certificaciones de calidad son uno de los requisitos para ingresar al régimen de promoción de la Ley de Software lo cual es un incentivo no menor para las empresas del sector, dados los beneficios que obtienen, tales como: descuentos de hasta el 60% del impuesto a las ganancias y acceso a bonos para pagar impuestos nacionales [1]–[4]. Sumado a esto, las certificaciones permiten acceder a mercados que exigen normas internacionales de calidad y aumentar sus oportunidades de negocios a través de la confianza que genera en sus clientes en relación a la seguridad y calidad de sus productos. [5].

Diversos estudios existen sobre los fracasos en la industria del software, tal es el caso del estudio presentado por la consultora internacional Standish Group, quien cita que tan solo el 32% de los proyectos de desarrollo de software, se pueden considerar exitosos, el 44% se entregaron fuera de plazo, excedieron

su presupuesto y no cubrieron la totalidad de las características y funcionalidad pactada, y el 24% de los proyectos fueron cancelados. En este mismo estudio se puntualiza que el principal factor para el fracaso de un proyecto de desarrollo de software radica en la mala calidad de los requisitos y la definición poco clara de los mismos [6].

Para obtener calidad en los productos de software hay que implementar prácticas de calidad en el proceso de desarrollo [7]–[10]. La base de dicho proceso, son los requisitos ya que determinan la calidad del producto final, razón por la cual la Ingeniería de Requisitos (IR) es una etapa crucial dentro del proceso de desarrollo de software (SW) [11], dentro de esta etapa una de las dificultades que se presenta es la especificación clara del comportamiento que se espera del software, ya que se tiene muchas veces sólo una descripción textual del mismo.

La elicitación de requisitos, subactividad de la IR tiene un alto impacto en el diseño y en las fases posteriores del ciclo de vida de un producto software. Si esta se realiza correctamente contribuye a reducir cambios y errores en los requisitos. Esta etapa se considera como la primera en el proceso de abstracción, comprensión y concepción de las funcionalidades que un producto software debe cumplir. [12].

El objetivo de la Ingeniería de Requisitos es desarrollar una especificación completa, consistente y no ambigua de los requisitos, la cual servirá como base para acuerdos comunes entre todas las partes involucradas y en dónde se describen las funciones que realizará el sistema [13].Tiene cuatro etapas fundamentales la elicitación, el análisis, la especificación y la validación. Durante ellas constantemente se requiere colaborar, coordinar y comunicar con los distintos stakeholders, entre sí, pero manteniendo distintas perspectivas y puntos de vista lo que plantea grandes retos de comunicación. [14]. Diversos problemas se presentan durante esta etapa, los cuales se deben a la carencia de un proceso adecuado de definición y entendimiento del problema y a la definición poco clara de las necesidades del cliente [15]. Pese a todos los estudios realizados, aún hoy los requisitos incompletos siguen siendo uno de los principales factores de fracaso de los proyectos [16]. Según Yang el 56% de los errores encontrados en el software son generados en la fase de requisitos y el porcentaje total del esfuerzo del proyecto que se destina en esta fase no llega al 20%, independientemente del modelo de desarrollo utilizado. [17].

Para intentar minimizar el costo y maximizar la calidad final del producto software es necesario estabilizar la especificación de requisitos [18], por lo anterior existen un conjunto de técnicas para evaluación y selección de los requisitos que buscan disminuir los costos y tiempos asociados a los mismos [19]–[21].

Dada la naturaleza cambiante de los requisitos, es necesario gestionarlos durante todo el desarrollo del software y considerar que la producción de los requisitos se caracteriza por ser sustancialmente interactiva e iterativa [22].

Uno de los mayores inconvenientes que se presenta dentro de esta etapa es la brecha de interpretación que se produce entre el Ingeniero de requisitos y el cliente lo que lleva a posibles errores de traducción de necesidades por parte del Ingeniero y a una falta de comprensión de los requisitos iniciales y una inhabilidad para gestionar los cambios durante el desarrollo. Por tal motivo es importante realizar una buena traducción de los requisitos y para esto se utilizan una serie de técnicas de elicitación a fin de reflejar las necesidades del usuario para ser integradas al software [23], [24]. Los errores que se presentan se podrían evitar si se hace uso de buenas prácticas en el marco de la ingeniería de requisitos. [25. Las técnicas de modelado o representación utilizadas normalmente en este marco no definen un proceso para tratar la complejidad contenida en el discurso de los interesados. Esto dota al proceso de análisis de un grado tal de imprecisión que hace que sea difícil llevar a cabo en forma efectiva esta actividad, dependiendo mucho de la experiencia de los analistas para seleccionar la técnica que mejor encaja en cada situación. Aunque muchos autores han desarrollado técnicas de modelado, al no contar con ninguna guía, la mayoría de los ingenieros prefieren utilizar las técnicas tradicionales. [18]. Pero al ser la IR una actividad que tiene un alto impacto

en el éxito o fracaso de los proyectos de desarrollo de software, es necesario implementar estrategias que garanticen el éxito de esta fase, entre ellos que apunten a aportar mejoras a la metodología, las técnicas, las herramientas utilizadas, así como también incrementar el compromiso y la participación de todos los involucrados, una tarea altamente compleja, por lo que es necesario utilizar técnicas específicas en la gestión del conocimiento a fin de mejorar la calidad de los requisitos, obtenidos en esta fase.

La ingeniería de conocimiento es una disciplina que es parte de la Inteligencia Artificial, cuyo fin es el diseño y desarrollo de sistemas expertos. Un Sistema Experto (SE), es un software que imita el comportamiento de un experto humano en la solución de un problema. Pueden almacenar conocimientos de expertos para un campo determinado y solucionar un problema mediante deducción lógica de conclusiones.

Diversos estudios avalan la existencia de problemas para satisfacer la necesidad de estructurar y categorizar la masa de información proveniente del proceso de educción de requisitos a los efectos de facilitar la comprensión del problema manifestado por el usuario. En otros términos, formalizar los requisitos mediante técnicas de representación del conocimiento [18].

Para dar solución a los problemas planteados se han llevado adelante estudios donde se analiza la aplicación de la Ingeniería del Conocimiento y de la gestión del mismo, en el campo de desarrollo de software. Ésta ha sido de gran utilidad dada la naturaleza misma de la Ingeniería de Software como una disciplina intensa en conocimiento [26]. Se ha analizado también el grado de concordancia entre los atributos de calidad de los conocimientos y de los requisitos, se concluye que ambos son coincidentes. Los mismos son exactitud, consistencia, certidumbre y fácil de usar [27].

Entre las técnicas más utilizadas para la representación de conocimiento se encuentran los mapas conceptuales, los cuales representan el conocimiento conceptual de un dominio. Los mismos se basan en la utilización de grafos que representan conceptos, objetos y relaciones entre ellos [28]. El eliminar ambigüedades y omisiones en los requisitos contribuye directamente a mejorar la calidad de los mismos y como consecuencia obtener un software de mayor calidad. La selección de esta técnica se debe a que los mapas conceptuales confeccionados por expertos tienden a exhibir altos niveles de acuerdo [29], [30]. Además, la utilización de Mapas Conceptuales tiene como fortaleza la generación de modelos de conocimiento [31].

En el proceso de Ingeniería de Conocimiento involucrando un experto, un investigador puede proyectar el software en una pantalla y trabajar junto al experto en la confección del diagrama, combinando de esta manera la Ingeniería de Conocimiento con la representación del conocimiento. Los enlaces pueden conectar también a otros mapas conceptuales, formando así un modelo de conocimiento. De esta manera los mapas conceptuales pueden constituir modelos de conocimiento estructurado y servir como repositorios de conocimiento de expertos. También pueden servir como interfaces para sistemas inteligentes para la construcción de herramientas de soporte o entrenamiento. Adicionalmente se explorarán técnicas alternativas.

Considerando este tipo de enfoques basados en inteligencia artificial, orientados a propiciar la calidad del proceso de elicitación de requerimientos, cabe destacar el trabajo de [32], cuyo objetivo es obtener una mejor representación de la información, partiendo de la automatización del proceso de generación de un mapa conceptual. De esta manera, se provee un mejor soporte para la interpretación y la toma de decisiones a partir de datos recolectados en el desarrollo del proyecto. Esta estrategia constituye una buena herramienta de gestión del conocimiento, ya que los conceptos pueden ser capturados o consultados, permitiendo el descubrimiento automático de conexiones implícitas para generar nuevos mapas. Los ingenieros del conocimiento extraen el conocimiento de los expertos humanos en una determinada área, y codifican dicho conocimiento de manera que pueda ser procesado por un sistema software [18].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

2.1. Relevamiento del estado de la industria en cuanto a IR

Se trabajará mediante estudios de casos a fin de hacer un análisis descriptivo de la situación actual de la ingeniería de requisitos en las empresas de software. Como primer parámetro de selección de empresas para la muestra del estudio se seleccionarán firmas que desarrollan software a medida a fin de que tengan la etapa de toma de requisitos, que cuenten con una persona o departamento que cumpla la función de ingeniero de requisitos y que tengan al menos diez empleados a fin de asegurar una estructura de funciones de desarrollo, entre otros criterios. finalmente se tendrá en cuenta que las firmas tengan al menos cinco años de funcionamiento en el mercado a fin de poder hacer un análisis de la trayectoria de las mismas en cuanto a los requisitos y los problemas y aprendizajes logrados en este aspecto. Se les realizará entrevistas en profundidad sobre las dimensiones abarcadas dentro del aspecto teórico a fin de identificar su comportamiento en cuanto a los requisitos y las consecuencias de ellos en sus resultados en el desarrollo de software.

2.2. Metodologías basadas en Conocimiento para la educción de requisitos

Las entrevistas permitirán detectar patrones comunes y distintivos de cada caso en cuanto al tratamiento y seguimiento de los requisitos. Se diagramarán los distintos procesos existentes en los casos a fin de determinar que técnicas de la ingeniería de conocimiento están presentes y cuáles no, a fin de detectar falencias y fortalezas en los diferentes procesos hallados.

3. RESULTADOS ESPERADOS

Se esperan detectar patrones comunes y distintivos de cada caso de estudio en cuanto al tratamiento y seguimiento de los requisitos.

Además, se pretende desarrollar una herramienta informática para el soporte integral de la ingeniería de requisitos basada en la Ingeniería de conocimiento, que será implementada en una empresa de la muestra a fin de determinar si produce mejora en tratamientos de requisitos y por lo tanto eleva la calidad del software desarrollado por la firma.

4. FORMACIÓN DE RRHH

El impacto sobre las capacidades institucionales estará garantizado por el contacto con los grupos de investigación del país y el exterior más consolidados.

El proyecto permitirá dar continuidad a la línea de investigación en la temática calidad de software iniciada en el año 2013 y contribuirá a la formación inicial como investigadoras de la Ing. Elena Ríos y la Lic. Lourdes Pralong.

Además, participan en el proyecto dos becarios alumnos de la Licenciatura en Sistemas de Información que inician su formación en la investigación.

5. BIBLIOGRAFÍA

[1] F. Barletta, M. Pereira, V. Robert, G. Yoguel, Argentina: Dinámica reciente del sector de software y servicios informáticos. 2013.

[2] R. L. Blanc, L. Lepratte, D. R. Hegglin, R. A. Pietroboni, Empresas de software en Entre Ríos: innovación, exportación, calidad y empleo, *XIX Reunión Anual Red Pymes Mercosur* 2014.

[3] F. Barletta, M. Pereira, and G. Yoguel, "Impacto de la Política de Apoyo a la Industria de Software y Servicios Informáticos," 2014.

[4] A. J. Pérez Puletti, "El sector de software y servicios informáticos de Argentina entre 2000 y 2012," 2014.

[5] R. Sosa Zitto, R. Blanc, L. Pralong, C. Álvarez, and S. Galáz, "El desafío de producir software de calidad aplicando prácticas de CMMI para las pymes de Concepción del Uruguay, Entre Ríos," in CoNaIISI 2013.

[6] T. Clancy, "The standish group chaos report," 2014.

[7] B. A. Nuseibeh and S. M. Easterbrook, "Requirements Engineering: A Roadmap," in ICSE 2000 Proceedings of the Conference on The Future of Software Engineering, 2010.

[8] A. F. Aguirre, C. Pardo, W. L. Pantoja, M. F. Mejía, and F. J. Pino, "Reporte de Experiencias de la Aplicación de Competisoft en Cinco Mipymes colombianas," Rev. EIA, vol. 7. 2013.

[9] G. Pantaleo, Calidad en el Desarrollo de Software. Alfaomega Grupo Editor, 2011.

[10] R. Sosa Zitto and R. Blanc, “Buenas prácticas de Scrum para alcanzar niveles de calidad en pymes de desarrollo de software,” in XVI Workshop de Investigadores en Ciencias de la Computación, 2014, pp. 490–494.
[11] C. M. Zapata, G. L. Giraldo, and J. E. Mesa, “Una propuesta de metaontología para la educción de requisitos,” Ingeniare. Rev. Chil. Ing., vol. 18, no. 26–37, 2010.
[12] M. A. C. Díaz and J. L. Carrillo-Medina, “Desarrollo de un Sistema Experto para la Especificación de Requerimientos,” Univ. Las Fuerzas armadas extensión Latacunga, vol. 8, p. 6, 2014.
[13] E. Kheirkhah and A. Deraman, “Important factors in selecting requirements engineering techniques,” in *Proceedings - International Symposium on Information Technology 2008, ITSim*, 2008, vol. 3.
[14] J. C. Ramírez Leal, W. J. Giraldo Orozco, and R. Anaya Hernández, “Una propuesta metodológica para mejorar la comunicación en IR,” Rev. EIA, vol. 13, no. 26, 2017.
[15] L. Londoño, R. Anaya, and M. Tabares, “Análisis de la Ingeniería de Requisitos Orientada por Aspectos según la Industria del Software,” Rev. EIA, no. c, pp. 43–52, 2008.
[16] M. Cristiá, “Introducción a la Ingeniería de Requerimientos,” 2011.
[17] Y. Yang, M. He, M. Li, Q. Wang, and B. Boehm, “Phase distribution of software development effort,” in Proceedings of the Second ACM-IEEE international symposium on Empirical software engineering and measurement, 2008, pp. 61–69.
[18] P. Pytel et al., “Propuesta de Aplicación de Técnicas de Representación de Conocimiento en el Análisis de Requisitos Software,”. Actas 1er Seminario Argentina-Brasil de TIC, 2011.
[19] R. Belschner, “Evaluation of Real-Time Requirements by Simulation Based Analysis,” in IFAC Proceedings Volumes, 1996.
[20] J. Karlsson and K. Ryan, “A Cost–Value Approach for Prioritizing Requirements,” IEEE Softw., vol. 14, no. 5, pp. 67–74, 1997.
[21] Ho-Won Jung, “Optimizing value and cost in requirements analysis,” IEEE 2002.
[22] P. (Pericles) Loucopoulos, V. Karakostas, and B. Καρακώστας, System requirements engineering. McGraw-Hill Book Co, 1995.
[23] V. Ambriola and V. Gervasi, “An environment for cooperative construction of natural-language requirement bases,” in Proceedings 8th Conference on Software Engineering Environments, 2002, pp. 124–130.
[24] J. Misra, “Terminological inconsistency analysis of natural language requirements,” Inf. Softw. Technol. 2016.
[25] L. E. P. Valencia, A. T. Lazo, and L. C. Benjumea, “La relación existente entre la calidad del software y el uso de modelos internacionales vs el uso de modelos autóctonos: el caso de las pymes en el eje cafetero–Colombia,” Informática, vol. 1, no. 1, 2013.
[26] J. Capote, C. J. Llanten Astaiza, C. J. Pardo Calvache, A. de J. González Ramírez, and C. A. Collazos, “Gestión del conocimiento como apoyo para la mejora de procesos software en las micro, pequeñas y medianas empresas,” Rev. Ing. e Investig., vol. 28. 2008.
[27] C. Hinojosa, G. Raura, and E. R. Fonseca, “El Aporte de la Gestión del Conocimiento a la Ingeniería de Requisitos,” Univ. Las Fuerzas armadas extensión Latacunga, vol. 8, p. 102, 2014.
[28] A. Sebastián and G. D. S. Hadad, “Mejoras a un modelo léxico mediante mapas conceptuales,” 2015.
[29] S. E. Gordon, “Implications of Cognitive Theory for Knowledge Acquisition,” in The Psychology of Expertise, New York, NY: Springer New York, 1992, pp. 99–120.
[30] K. Hoffman, RR and Coffey, JW and Ford, “A case study in the research paradigm of human-centered computing: Local expertise in weather forecasting,” 2000.
[31] R. R. Hoffman and G. Lintern, “Eliciting and Representing the Knowledge of Experts,” in The Cambridge Handbook of Expertise and Expert Performance, 2006.
[32] W. M. Wang, C. F. Cheung, W. B. Lee, and S. K. Kwok, “Self-associated concept mapping for representation, elicitation and inference of knowledge,” Knowledge-Based Syst., vol. 21, no. 1, pp. 52–61, 2008.

Un proceso para Desarrollo Dirigido por Modelos en entornos Agiles

Alberto Cortez[1,2], Carlos Martinez[1], Claudia Naveda[1,2],
Javier Caballero[1], Matías Luna[1], Alejandro Vazquez[1].

UTN Facultad Regional Mendoza, Ingeniería en Sistemas de Información,
[1]Laboratorio de Auditoria y Seguridad de TIC,
[2]Universidad de Mendoza, Instituto de Informática

{cortezalberto, caballerojavier13, claudialaboral, mtsluna }@gmail.com ,
avazquez@frm.utn.edu.ar, carlos.martinez@ frm.utn.edu.ar

RESUMEN

La industria del software integra, en su proceso de desarrollo, las metodologías ágiles como un estándar, porque brindan valor agregado a los usuarios en el menor tiempo posible y además se adapta a los cambios del negocio en forma oportuna. Sin embargo, muchas de estas técnicas carecen de definiciones de sus procesos completamente. Se observa también que no hay herramientas que ofrezcan soporte sobre a los mismos.

El desarrollo de software dirigido por modelos permite mejorar esta situación, incorporando al proceso de desarrollo de software la abstracción y el formalismo necesario, para automatizar y optimizar las tareas más críticas del proceso de desarrollo.

La integración de ambas metodologías produjo un nuevo enfoque denominado Desarrollo de software dirigido por modelos ágil, (AMDD, Agile Model Driven Development), que permite acelerar el desarrollo a partir una perspectiva ágil.

La línea de investigación propone en el contexto de AMDD, incorporar un proceso metodológico y crear herramientas con MDD (Model-Driven Development) para que sea aplicado a un proyecto ágil.

El presente trabajo se diferencia de otros, porque efectúa procesos de trazabilidad de la evolución de los artefactos, durante el proceso de construcción de proyectos de software. Se emplean aportes de MDD para: generación automática de código, ingeniería inversa y transformaciones de modelos, en conjunción con metodologías ágiles.

Palabras clave: AMDD, Procesos. Desarrollo Ágil, MDD, Transformaciones de modelos, Ingeniería inversa.

CONTEXTO

La línea de investigación, de reciente formación, de "Desarrollo Dirigido por Modelos en entornos Agiles" que se presenta en este trabajo, se desarrolla en el marco del proyecto Interinstitucional conformado por el Instituto de Informática de la Universidad de Mendoza y el GRUPO AuSegTIC (Grupo de Auditoría y Seguridad de TIC) de reciente formación, de la Facultad Regional Mendoza de la Universidad Tecnológica Nacional.
La misma es parte de un proceso de incentivación para el desarrollo de actividades I/D, que se originan en el marco de este acuerdo.

1. INTRODUCCIÓN

Los métodos ágiles son procesos más rápidos de desarrollo de software, que surgieron como respuesta a las limitaciones que producen los métodos de trabajo: tanto ausencia como

exceso en la cantidad de los mismos [1]. Ellos imponen un enfoque iterativo para desarrollar sistemas de manera incremental. Las metodologías ágiles incluyen Programación Extrema [2], Scrum [3], Crystal [4], etc. Cada uno de ellos comparte principios y valores comunes definidos en el Manifiesto Ágil [5].

Los métodos de desarrollo guiado por modelos (por sus siglas en inglés MDD) [6] surgieron como un nuevo paradigma de desarrollo de software, donde los modelos juegan un papel fundamental. Se emplean para especificar un sistema y generar automáticamente su código. MDD se creó con el objetivo de elevar el nivel de abstracción, y además aumentar la automatización en la generación de código según propone Frankel en [7]. De esta manera, es que MDD pretende mejorar la productividad, la portabilidad, la interoperabilidad y el mantenimiento de los sistemas según Kleppe et al. en [8]. La integración de Agile y MDD (Agile MDD) es de creciente interés por muchas razones. En primer lugar, los beneficios que proporcionan Agile y MDD pueden llevar a un mejor desarrollo de software afirman Vijayasarathy et al. [9]. En segundo lugar, el MDD se puede considerar como un proceso de desarrollo lento, lo que puede ser una barrera para su adopción en la industria, y agregarle agilidad facilitará su adopción. Existen diferencias fundamentales y conflictos entre metodologías ágiles y MDD que hacen que la integración sea bastante desafiante. En algunos trabajos presentados en la literatura, se observan tanto procesos que integran ambas metodologías como la ausencia de herramientas adecuadas para trabajar en dicha integración.

La idea básica detrás de ambos modelos, MDE y Metodologías ágiles, es crear sistemas que puedan responder rápidamente a cambios frecuentes, se proponen diferentes enfoques para resolver los requisitos mencionados; la agilidad se enfoca en aspectos metodológicos que conciernen a un producto individual, mientras que MDE centra su estudio en un aspecto arquitectónico definido por su variante específica MDA (Model Driven Arquitectura) [10] que separa las características del sistema de su implementación en una plataforma técnica. Advirtiendo la importancia de metodologías ágiles y MDA en el desarrollo de software de sistemas, muchos trabajos se centraron en combinar estas áreas por lo que en esta sección se presentan algunos trabajos previos.

En su propuesta Essebaa et. al [11] presentan el proceso V Life Cycle y lo usan para definir los pasos del desarrollo de sistemas. Proponen una combinación de MDA y Scrum + V Life Cycle. Modelo V significa modelo de Verificación y Validación. Al igual que el modelo de cascada, el ciclo de vida V es un camino secuencial de ejecución de procesos. Cada fase debe ser completada antes de comenzar la siguiente; basado en el documento de requisitos que contiene especificaciones del sistema, el equipo de desarrolladores comienza el trabajo de diseño y luego comienza la real implementación del código y el equipo de testing comienza con la planificación, escritura de casos de pruebas y pruebas de scripting. Ambas actividades comienzan en paralelo

En su artículo [12] M.B.Nakicenovic presenta un proceso AMDD que fue desarrollado considerando prácticas livianas y ágiles, el documento tiene como objetivo proporcionar un enfoque que demuestre que MDD y metodologías ágiles pueden trabajar juntos, aprovechando los beneficios de cada dominio. El enfoque se aplica tanto en ingeniería directa como inversa para responder a dos cuestiones: la aceleración del proceso de reingeniería de la solución MDD, cómo beneficiarse de la agilidad y liviandad al producir una solución MDD en un período de tiempo corto. El artículo describe un enfoque que combina

MDD y metodologías ágiles basado en su versatilidad. La implementación del enfoque se realizó en el proyecto Market Server Capabilities (MSC) propuesto por la empresa SunGard. Los autores F.P. Basso et al. presentan una propuesta que combina los enfoques de MDA y metodologías ágiles en el contexto del Prototipado de Aplicación Rápida (RAP por sus siglas en inglés) [13]. RAP permite la validación de los requerimientos de software, antes de las pruebas de aceptación, con el fin de obtener una respuesta rápida de los clientes. Este enfoque tiene en cuenta los principios de metodologías ágiles en el contexto de MDA y basándose en la metodología RAP para generar modelos y Front End, fundados en el patrón MVC, la implementación del enfoque se apoyó en metodología scrum y MDA para la generación de sistemas de información web. Los autores pretenden asegurar varios beneficios dentro de este enfoque; mejor organización del código fuente, simplicidad en los cambios del código fuente, diseño simple y rápido de modelos, y beneficios de la reutilización de diseño de modelos. Pero este enfoque no puede ser generalizado a todo tipo de sistemas de software, de hecho, se dedicó para desarrollar Sistemas de Información Web. Y tampoco los autores de este enfoque detallan cómo integrar MDA y Scrum, como por ejemplo no proponen dónde usar cada nivel de MDA en la metodología Scrum.

V.Kulkarni et al. discuten en su trabajo [14] por qué la metodología ágil no se puede utilizar con MDE, entonces proponen una modificación a hacer sobre metodologías ágiles de manera de combinarlas con MDE. En consecuencia este documento describe un nuevo proceso de desarrollo de software que combina Scrum y MDE, los autores propusieron el uso de Meta-Sprints que se ejecuten en paralelo a Sprints para validar los modelos, y sugieren dos o tres meses como escalas de tiempo para meta-sprints, en los cuales los clientes deben proporcionar retroalimentación sobre modelos y prototipos, algo que es opuesto a los principios ágiles; de hecho, la agilidad recomienda que la retroalimentación de los clientes sea en un período inferior al que se propone en este enfoque.

H. Alfraihi en su artículo [15] analiza el desafío de combinar Ágil y MDD, el paper describe un enfoque que pretende aumentar la adaptabilidad de estos dominios proponiendo un framework que aligere tanto Agilidad como MDD, este enfoque propone recomendaciones, pautas, y procedimientos para poder usar Agile MDD en la práctica. Se observa que incluso aunque este enfoque propone algunas prácticas para implementar Agile MDD, no tiene en cuenta la arquitectura de MDD, MDA y cómo beneficiarse de los diferentes niveles de abstracción para producir sistemas de software sostenibles.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

En base a lo expresado previamente, en esta línea de investigación se trabaja en el estudio, definición y uso de un proceso para la integración de Ágil con desarrollo de Software dirigido por modelo. Será necesario para el desarrollo del trabajo, el estudio y análisis de los estándares y herramientas que participan en MDD y la metodología de Scrum.

En particular, se proponen dos instancias para su definición e implementación: (i) el desarrollo de herramientas que cumplan los estándares de MDD; y (ii) definición de un proceso que integre en forma eficaz ambas metodologías valiéndose de las herramientas creadas y generando valor agregado a su implementación en la producción de software.

3. RESULTADOS OBTENIDOS Y OBJETIVOS

La línea de investigación abordada se definieron los siguientes objetivos:

Objetivos Generales:

Contribuir en la incorporación de un proceso formal, herramientas de producción automática de prototipos y de gestión administrativas al manejo de evolución y trazabilidad de modelos que acompañen a ciclo de vida de Scrum.

Objetivos Específicos:

- Crear una Lenguaje Específico de Dominio [16].

- Formalizar el DSL con la aplicación de reglas de buena formación a través de OCL.

- Diseñar una sintaxis gráfica concreta del DSL, amigable al usuario a través de Sirius [17].

- Crear un repositorio de plantillas para realizar transformaciones de modelo a texto a través de Acceleo [18].

- Crear transformaciones de ingeniería inversa con Modisco [19], para sincronizar los programas fuentes que desarrollan con las distintas versiones del modelo.

- Diseñar un software Web para llevar la auditoría y trazabilidad de versiones resultantes.

El proceso que se presenta en el siguiente trabajo define las siguientes etapas y actividades que se pueden llevar en la conjugación de MMD y Ágil para producir software de calidad y a menor costo.

Proceso:

Fase 1: Determinar las historias de usuarios prioritarias y generar el primer Sprint.

Fase 2: Diseñar el Prototipo inicial del modelo es una instancia colaborativa con el equipo de Scrum, el mismo brindará una idea inicial de las entidades que se reflejan en las historias de usuario como así también una idea general de la arquitectura del sistema.

Fase 3: Realizar la generación automática de código del modelo Inicial, para las tecnologías de Front End, Back End y Persistencia elegida. Por ejemplo: ReactJS, Node y Mongo Db.

Fase 4: Realizar el despliegue del prototipo al Repositorio de Modelos con su correspondiente Versión y la distribución de tareas a cada miembro de desarrollo.

Fase 5: El líder de proyecto utilizará el proceso de ingeniería inversa por cada miembro del grupo de desarrollo para sincronizar los cambios hechos en el código, para que de forma automática actualice el modelo inicial generando una nueva versión del mismo.

Fase 6: El equipo de desarrollo realizará el despliegue de su desarrollo al servidor de versionado, registrando el número de versión.

Fase 7: El líder de proyecto integrará el repositorio de modelos y el servidor de versionado para que las versiones de los modelos estén sincronizadas con las versiones del servidor de despliegue.

Estas fases del proceso se deberán repetir para iterar sobre el sprint en cuestión hasta que se produzca el primer entregable con la aceptación del cliente.

En la figura 1, se puede observar el proceso propuesto con los roles involucrados modelado con BPMN [20].

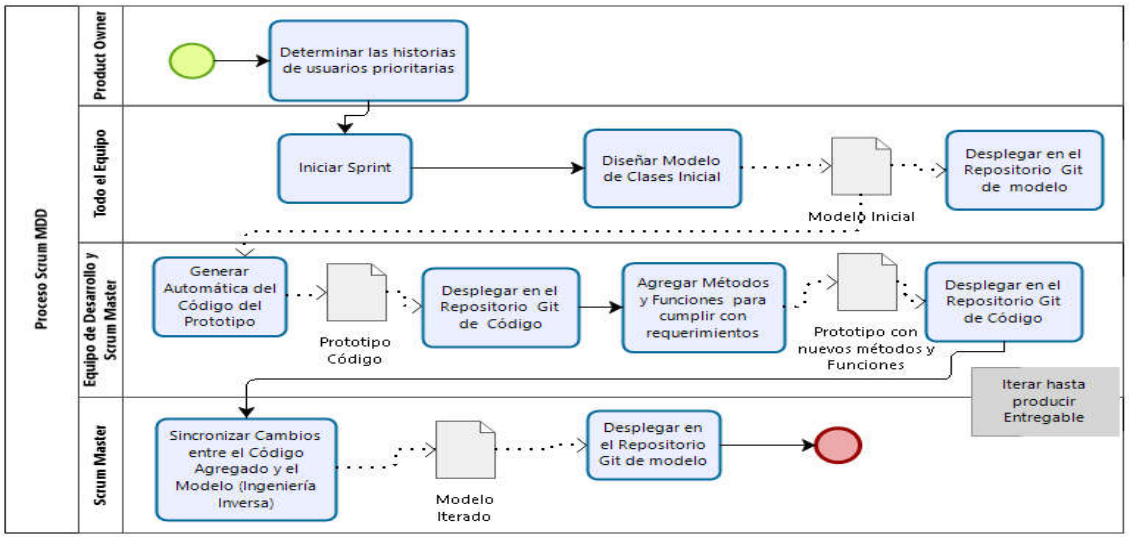


Fig.1: Proceso propuesto.

4. FORMACION DE RECURSOS HUMANOS

Durante el desarrollo de esta línea de investigación, integrantes del grupo de trabajo están actualmente trabajando es su plan de Tesis de Doctorado, tesinas de grado en ambas universidades. También, se están formando ayudantes de segunda en las asignaturas de la cátedra de Desarrollo de Software Dirigido por Modelos y de Diseño de Sistemas. Consecuentemente, se están elaborando seminarios abiertos de formación general en relación a temas, técnicas y tecnologías incluidos en esta línea de investigación, para alumnos avanzados y docentes en la carrera de Ingeniería en Sistemas y afines.

5. BIBIOGRAFÍA

[1] M. Fowler. (2005, December,13). The new methodology [Internet] Available: http://martinfowler.com/articles/new-Methodology.html.

[2] K. Beck and M. Fowler, Planning Extreme Programming. The XP series. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA 2000.

[3] K. Schwaber, and M. Beedle. Agile Software Development with Scrum. Agile Software Development. Prentice Hall, 2002.

[4] A. Cockburn, Agile Software Development. Agile Software Development. Addison-Wesley, 2002.

[5] K. Beck .Manifesto for agile software development, 2001. [Online]: Available: https://agilemanifesto.org/

[6] OMG, "The pragmatics of model-driven development", IEEE Softw., vol. 20, no. 5, pp. 19 25, 2003.

[7] D. Frankel. Model Driven Architecture: Applying MDA to Enterprise Computing, volume 25. Hoboken, New Jersey John Wiley & Sons, 2003.

[8] A.Kleppe, J. Warmer, and W. Bast, MDA Explained: The Model Driven Architecture: Practice and Promise. The Addison-Wesley object technology series. Primera Edición. Reading, Massachusetts Addison-Wesley, 2003.

[9] L. Vijayasarathy and D. Turk, "Agile software development: A survey of early adopters". Journal of Information Technology Management, 19(2):1–8, June 2008.

[10] M. Siegel, MDA (2014, June,6). [Online]: Available: https://www.omg.org/mda/

[11] I. Essebaa and S. Chantit. "Model Driven Architecture and Agile Methodologies", Proceedings of the Federated Conference on Computer Science and Information Systems pp. 939–948, ISSN 2300-5963 ACSIS, Vol. 15, 2018.

[12] M. Nakićenović, "An agile driven architecture modernization to a model-driven development solution," International Journal on Advances in Software, vol. 5, no. 3,4, 2012.

[13] F. Basso, R. Pillat, F. Roos-Frantz, and R. Z. Frantz, "Combining mde and scrum on the rapid prototyping of web information systems,", Int. J. Web Engineering and Technology, vol. 10, no. 3, October 2015.

[14] V. Kulkarni, S. Barat, and U. Ramteerthkar, "Early experience with agilemethodology in a model-driven

approach," in Model Driven Engineering Languages and Systems, 14th International Conference, MODELS 2011, Wellington, New Zealand, October 16-21, 2011. Proceedings, 2011, pp.578–590.

[15] H. Alfraihi, "Towards improving agility in model-driven development," in Joint Proceedings of the Doctoral Symposium and Projects Showcase Held as Part of STAF 2016 co-located with Software Technologies: Applications and Foundations (STAF 2016), Vienna, Austria, July 4-7, 2016., 2016, pp. 2–10.

[16] Eclipse Modelling Framework, último acceso: febrero 2019. [Online]: Available: https://www.eclipse.org/modeling/emf/

[17] Sirius, último acceso: febrero 2019. [Online]: Available: https://www.eclipse.org/sirius/

[18] Acceleo, último acceso: febrero 2019. [Online]: Available: https://www.eclipse.org/acceleo

[19] Modisco, último acceso: febrero 2019. [Online]: Available: https://www.eclipse.org/MoDisco/

[20] BPMN, último acceso: febrero 2019. [Online]: Available: https://www.omg.org/bpmn/

.

Usabilidad en Simuladores Web de Redes de Sensores Inalámbricos

Diego Alberto Godoy, Hernán Bareiro, Eduardo O. Sosa, Emilce Stoffel[a], Gabriela Barros[b]

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

{diegodoy; hbareiro; eduardo.sosa; @citic.ugd.edu.ar}, [a]emiest89@gmail.com, [b]gabubarros07@gmail.com

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 un estudio sobre la usabilidad de dos nuevos simuladores Basados en la Web, WEBShawn y NS2web. Además se propone una metodología de trabajo para realizar este estudio.

Palabras claves: Simulación; WSN; Usabilidad; Proyecto de Desarrollo de Software; Simulación Basada en la Web.

Contexto

De acuerdo a la norma ISO 25000, usabilidad es la Capacidad del producto software para ser entendido, aprendido, usado y resultar atractivo para el usuario, cuando se usa bajo determinadas condiciones. El presente trabajo se desarrolla en el Centro de Investigación en Tecnologías de la Información y Comunicaciones (CITIC), en el contexto del proyecto "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 con el Código IP A07003 de la Secretaría de Investigación y Desarrollo de la Universidad Gastón Dachary (UGD).

1. Introducción

Una red de sensores inalámbricos (WSN) es una red de interconexión inalámbrica que consiste en dispositivos distribuidos espaciados y autónomos utilizando sensores para monitorear condiciones físicas o ambientales.

La simulación aplicada a las redes de sensores inalámbricos es un componente clave para simplificar y ayudar en diferentes áreas de este tipo de redes. Puesto que tiene positivas incidencias al momento de desarrollar aplicaciones que se adapten a las especiales características con que cuentan. Como ser por ejemplo, la posibilidad de que en una misma red se encuentren interconectados cientos o miles de nodos realizando distintos procesos cada uno.

Ya sea que se utilice la simulación con fines de implementación en una red real, virtual o en trabajos de investigación académicos. Se buscan simuladores de fácil instalación, configuración, accesibilidad y potentes para simular posibles escenarios que incluyan millares de nodos sensores.

Al momento de poner en funcionamiento un sistema de simulación, se buscan herramientas portables o multiplataforma, que tengan una interfaz que abstraiga al usuario de la complejidad necesaria en su configuración

e instalación. Este tipo de herramientas evitan que los tiempos de un proyecto de simulación se trasladen en mayor medida a la instalación y puesta en marcha del simulador.

Una de las principales características de los sistemas Web es su portabilidad, así también como su facilidad de acceso en forma remota. Con el uso del modelo cliente-servidor al utilizar, un navegador Web como interfaz, se abstrae al usuario de toda la complejidad interna de un sistema. La complejidad en este caso queda completamente transparente del lado del servidor.

Existe un simulador Basado en la Web, denominado WEBShawn [1], construido sobre las bases de Shawn [2], que es un simulador de redes de sensores inalámbricos de interacción por línea de comandos. Al simulador Shawn se le realizaron adaptaciones en su funcionamiento principal para que se puedan realizar corridas de simulación y visualización de resultados a través de la Web, permitiendo cambiar la forma de interacción del usuario con el simulador; que originalmente fue desarrollado para simulaciones vía consola de comandos. El simulador WEBShawn [3] se encuentra desarrollado pero no se ha testeado aún a nivel de usuario. La experiencia de usuario es importante para conocer la usabilidad de la aplicación y descubrir las necesidades de adaptación que pueda requerir la misma [4].

Otro simulador Basado en la Web es NS2web [5] [1]. El mismo permite la ejecución remota de simulaciones con NS-2 para escenarios cableados e inalámbricos. El software también proporciona un kit de herramientas para analizar archivos de rastreo generados como resultado de la simulación. Esta aplicación permite al usuario centrarse en las simulaciones con NS-2, en lugar de gastar su tiempo en instalar el simulador o escribir scripts personalizados para su análisis [5].

Para evaluar la calidad de un sitio Web se tienen en cuenta varios atributos, entre ellos la usabilidad, que es considerado de suma importancia ya que indica la facilidad de uso que tiene un artefacto para un usuario según ciertas condiciones determinadas [6] [7].

La Organización Internacional para la Estandarización propone dos definiciones relativas a la usabilidad del software:

ISO/IEC 9126: "La *usabilidad* se refiere a la capacidad de un software de ser comprendido, aprendido, usado y ser atractivo para el usuario, en condiciones específicas de uso".

ISO/IEC 9241: "Usabilidad es la efectividad, eficiencia y satisfacción con la que un producto permite alcanzar objetivos específicos a usuarios específicos en un contexto de uso específico" [6].

De este modo, la usabilidad define la interacción del usuario con las aplicaciones, en el caso particular del presente trabajo se enfocará en la interacción del usuario con los simuladores de WSN Basados en la Web, centrándose en la facilidad de navegación y rendimiento de las tareas que llevan a la realización de las transacciones [4].

Para recolectar los datos de interacción del usuario con un producto de software y su posterior evaluación se requiere un método [8]. Los métodos de evaluación de usabilidad pueden ser clasificados por numerosos criterios, debido a que aún en la actualidad no existe un acuerdo unificado para clasificarlos y a que los diferentes autores e investigadores definieron sus propios métodos. Lo que sí se observa son entrecruzamientos con otras categorías y coincidencias en las mismas.

Los métodos de evaluación más utilizados son los métodos de inspección, empíricos y de indagación (Figura 1) [9] [10] [11] [12] [13] [14].

Figura 1. Métodos de evaluación de usabilidad

Evaluar la usabilidad de las aplicaciones Web, se ha convertido en un proceso crítico en los últimos años, ya que se trata de entrar a competir con calidad en el mundo de productos y servicios de Internet, el cual se encuentra saturado. Esta problemática motiva el objetivo del presente trabajo, que es presentar el estado del arte de los modelos y métodos de evaluación de usabilidad orientados a los simuladores Web de WSN. En este sentido no existe registro de que WEBShawn y NS2Web hayan sido testeadas con usuarios finales, por lo cual resulta necesario probarlas a fin de proponer mejoras en cuanto a la usabilidad.

2. Línea de Investigación

Para esta línea se han planteado los siguientes objetivos:

Como objetivo general se propone:

- Evaluar la usabilidad y las opciones de configuración de simuladores Web de WSN en base a la experiencia de usuarios de carreras TIC.

Como objetivos específicos se realizarán los siguientes:

- Indagar el estado del arte de simuladores Web de WSN, en particular WEBShawn y NS2web.
- Estudiar los modelos, normas y/o estándares de usabilidad.
- Seleccionar un subconjunto de criterios de usabilidad para evaluar y comparar las opciones de configuración de los simuladores Web de WSN.
- Realizar pruebas de los simuladores Web de WSN con usuarios de carreras TIC, para analizar su experiencia al utilizarlos.

3. Resultados

Como avances del proyecto se ha definido el método de trabajo el cual detalla el procedimiento y observaciones realizadas con los instrumentos de recolección de datos utilizados. En la Figura 2 se puede observar una guía orientativa de cómo llevar a cabo el método definido.

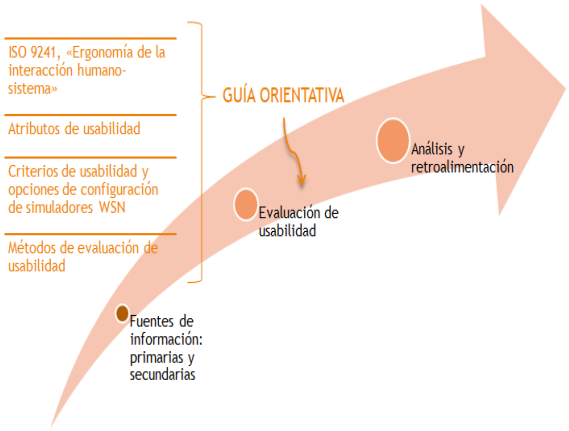


Figura 2 . Método de trabajo y procedimiento de evaluación de usabilidad.

Con este método de trabajo se deben detallar los participantes de las pruebas, los actores que participan de las pruebas y las actividades que se desarrollan. Se considera importante disponer de métodos y herramientas que contribuyan a comprobar fehacientemente el grado de usabilidad de los simuladores Web de WSN, promoviendo la participación del usuario en las pruebas. Para realizar las pruebas con usuarios se definieron dos tipos. Usuarios alumnos y usuarios profesionales.

Con los usuarios alumnos se realizaron dos jornadas presenciales en el laboratorio del campus de la UGD. Participaron nueve alumnos (Figura 3).

Con los usuarios profesionales se realizó una convocatoria a los profesionales del Departamento de Ingeniería y Ciencias de la Producción de la UGD. Se enviaron correos electrónicos en forma particular a distintos profesionales pidiendo colaboración. Utilizaron el simulador ocho profesionales.

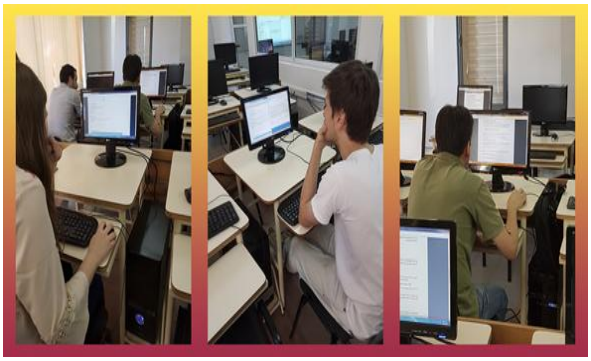


Figura 3. Alumnos utilizando el simulador

Se diseñó un instrumento de recolección (cuestionario) de información, con el propósito de determinar el grado de importancia que conceden a la usabilidad los profesionales que usan simuladores Web de WSN, como así también los usuarios/alumnos. Se utilizó el Laboratorio de Informática de la Universidad Gastón Dachary (UGD), sede Campus para realizar las experiencias con los simuladores y responder al cuestionario.

A los profesionales se les consulto sobre: Propiedades deseables de un simulador, Disponibilidad y reutilización, Rendimiento y escalabilidad, Soporte para la definición de experimentos y resultados mediante el uso de scripts de simulación, Soporte gráfico, de depuración y ejecución paso a paso.

A los alumnos se les consulto sobre: Usabilidad del sitio y del simulador, aspectos generales, identidad e información, estructura y navegación, rotulado, layout de página, capacidad para entender la información expuesta y facilidad de uso, control y retroalimentación y ayuda.

4. Formación de Recursos Humanos

El equipo de trabajo se encuentra formado por cuatro investigadores, 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; y ocho estudiantes en período de realización de trabajos finales de grado en de las carreras de Licenciatura en Sistemas de Información y de Ingeniería en Informática. Actualmente, el número trabajos finales de grado aprobado, es de ocho, y otras tres en proceso de desarrollo. Además de una tesis doctoral finalizada.

5. Bibliografía

[1] Barun Kumar Saha. NS2Web. [Online]. http://vlssit.iitkgp.ernet.in/

[2] Fischer, S., Pfisterer, D. and Fekete, S. P., "Shawn: The fast, highly customizable sensor network simulator," *Braunschweig University of Technology and University of Lubeck, Alemania, 2007. ISBN 1-4244-1231-5.*

[3] 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.

[4] J. NIELSEN, "Usability Engineering," *Morgan Kaufmann Publishers*, vol. 1 edition, September 23 1993. ISBN-10: 0125184069.

[5] B. K. SAHA, S. MISRA and M. S. OBAIDAT, "A Web-based Integrated Environment for Simulation and Analysis with NS-2," *IEEE Wireless Communications*, vol. 20, no. 4, pp. 109-115, August 2013.

[6] E. SERRANO MASCARAQUE, "Accesibilidad vs usabilidad web: evaluación y correlación. INVESTIGACIÓN BIBLIO TECOLÓGICA," vol. 23, no. 48, Mayo/ Agosto México. 2009. ISSN: 0187-358X.

[7] J. NIELSEN and M. N., TAHIR, *Usability de páginas de inicio: análisis de 50 sitios web*.: Pearson Educación, 2002.

[8] A. FERNÁNDEZ and E. INSFRAN, "Usability evaluation methods for the web: A systematic mapping study," *Journal Information and Software Technology*, Vol. 53 Issue 8, August, 2011.

[9] M., HEARST, M. IVORY, "The state of the art in automating usability evaluation of user interfaces. ," *Journal ACM Computing Surveys (CSUR)*, Vol. 33, No. 4, December 2001.

[10] A. FERNÁNDEZ MARTÍNEZ, "Usability Inspection Method for Model-driven Web Development Processes. ," Ph.D. THESIS. Valencia, Universitat Politècnica de València (UPV). 2012.

[11] J. BASTIEN, "Usability testing: a review of some methodological and technical aspects of the method," *International Journal of Medical Informatics*,. Vol. 79, Issue 10, October 2010.

[12] L. VAN WAES, "Thinking aloud as a method for testing the usability of websites: The influence of task variation on the evaluation of hypertext," *IEEE Transactions on professional communication*, Vol. 43, No. 3, 2000.

[13] E. KRAHMER and N. UMMELEN, "Thinking about thinking aloud: A comparison of two verbal protocols for usability testing," *IEEE Transactions on professional communication*, Vol. 47, No. 2, 2004.

[14] F. STEFANO and S. BORSCI, "Web usability evaluation with screen reader users: implementation of the partial concurrent thinking aloud technique," *Cognitive processing*, Vol. 11, Issue 3, 2010.

Innovación en

Educación en

Informática

Analíticas del aprendizaje en el uso de videos en el ámbito universitario

Alejandra Orellana; Rosa, Pósito

Departamento de Informática / FCEFN/ UNSJ/GATIE
Av. Ignacio de la Roza 590 (O) - San Juan- Argentina
Telefonos:4260353 -4260355
ale_ore@yahoo.com.ar; rosaposito@gmail.com

Resumen:

El solo hecho de utilizar un video y su sola visualización no garantiza que el alumno haya adquirido alguna competencia, por ello hay que completarlo con un análisis de interacciones virtuales. Las analíticas del aprendizaje emergieron inicialmente para estudiar datos a posteriori, describiendo resultados y aportando diagnósticos. Hoy en día han evolucionado pues ayudan a los docentes a identificar qué alumnos avanzan con dificultad, cuáles son las áreas más débiles, y así identificar factores que afectan el éxito de un aprendizaje más reflexivo. Para la implementación de la analítica del aprendizaje, es de vital importancia trabajar sobre entornos virtuales flexibles que permitan distribuir material audiovisual, no solo para su visualización, sino también para la recogida de datos. Los desempeños pueden ser visualizados mediante informes que ayudan a determinar el perfil del alumno, paneles de visualización de tendencias, o gráficos que permitan dar valor a las notas obtenidas y permitan fortalecer el trabajo docente a la hora de recopilar resultados sobre este proceso y sean de utilidad para la toma de decisiones sobre los recursos utilizados, las estrategias propuestas como así también sobre los modelos de enseñanza aplicados en el aula.

Palabras claves: analítica del aprendizaje, video, entornos virtuales

Contexto

El presente artículo se realiza en el marco del Proyecto de Investigación "*Los entornos virtuales flexibles para el desarrollo de competencias en el área de algoritmos y lenguajes de programación*" Código 21/E1113, Vigentes en el periodo 2018-2019, de la Universidad Nacional de San Juan. Argentina CICITCA. Resolución N°21/18-CS. Este proyecto forma parte del trabajo conjunto del Programa permanente de Investigación del Gabinete de Tecnología e Innovación Educativa (GATIE) y el Departamento de Informática. El Proyecto tiene el propósito de mejorar el proceso de enseñanza-aprendizaje de los alumnos de las carreras de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información de la Universidad Nacional de San Juan, mediante la investigación de herramientas, aplicaciones, entornos de trabajo formales e informales, que permitan comprender cómo están involucrados nuestros alumnos en su proceso de aprendizaje. Como así también evaluar el uso de dichas herramientas de modo tal que el docente, pueda generar nuevas estrategias didácticas pedagógicas en función de los datos recolectados durante el proceso de enseñanza-aprendizaje. Además, se tienen en cuenta teorías psicosociales, modelos pedagógicos que beneficien el desarrollo de competencias básicas para los alumnos de ambas carreras. Se aspira que los resultados de esta investigación permitan fortalecer el trabajo docente a la hora de recopilar resultados sobre este proceso y sean de utilidad para la toma de decisiones sobre los recursos utilizados, las estrategias propuestas como así también sobre los modelos de enseñanza aplicados en el aula.

El equipo de trabajo está formado por docentes investigadores del Departamento de Informática que trabajan también en el Programa Permanente de Investigación del Gabinete de Tecnología e Innovación Educativa, por sus siglas denominado GATIE.

En el presente artículo, se expone como línea de I/D la aplicación de las analíticas del aprendizaje en el uso de videos educativos diseñados para dispositivos móviles.

1. Introducción

Como señala (Bravo Ramos, 2000), "la introducción del vídeo en el aula puede producir modificaciones sustanciales en el escenario donde tiene lugar la docencia." La presentación del vídeo no debe verse como una forma de entretener a los alumnos. El vídeo debe tener un objetivo didáctico previamente formulado. El aprendizaje no se encuentra en función del medio, sino fundamentalmente en las estrategias y técnicas que se apliquen sobre él. También es muy relevante la actitud y la estrategia didáctica que el docente tenga al presentar el vídeo como un material didáctico, ya que a través de ello se deriva gran parte del aprendizaje en los alumnos.
Asociado al uso de los órganos sensoriales más utilizados por el ser humano, (Zankov, L. V., 1977) discípulo de la escuela de Vigotski postula que el órgano sensorial que controla el mayor ingreso de estímulos de entrada es la vista, con un 83%, luego le sigue el oído con un 11% y el kinestésico en menor grado. Llamamos a estos canales perceptivos.

El vídeo es uno de los medios audiovisuales por excelencia y ha demostrado tener un gran potencial para facilitar el aprendizaje activo y combinado. Diversos estudios han demostrado la capacidad del video para involucrar al alumno y activar el aprendizaje cognitivo y emocional.

Tener en cuenta los canales perceptivos, proporciona información clave para seleccionar diversos tipos de materiales didácticos en función del modo en que los alumnos puedan aprender mejor. Pero además de estos canales, es importante conocer sobre los estilos cognitivos o de aprendizaje, modelo propuesto por (Kolb, D., 1984) éstos, vendrían a ser una descripción de las formas de comportarse de los alumnos y que determinan la preferencia que tienen de aprender, es decir cómo utilizan los recursos cognitivos tales como la forma de percibir, el uso de la memoria, el procesamiento de los datos, etc., estos estilos cognitivos se pueden resumir en cuatro: teóricos, reflexivos, pragmáticos y activos. Aquí las analíticas del aprendizaje se relacionan con la combinación de los canales perceptivos y los estilos cognitivos en la forma en que los alumnos reciben, procesan y utilizan los datos y la información, de vital importancia para establecer itinerarios de aprendizaje más individualizados (Amo, D, Santiago, R., 2017).

El solo hecho de utilizar un video y su sola visualización no garantiza que el alumno haya adquirido alguna competencia, por ello hay que completarlo con un análisis de interacciones virtuales. Las introducciones de cuestionarios dentro de los videos difieren de otro tipo de evaluaciones en que se muestra directamente dentro del visor del video, por lo que los alumnos reflexionan sobre el material visto hasta el momento o buscan hacia atrás para revisar el material relacionado con el cuestionario, también pueden omitir el cuestionario, etc. La presencia de cuestionarios dentro de los videos puede influir en los comportamientos de visualización de los alumnos y se podría obtener ciertos patrones de navegación no lineales. Estos estudios comprueban que los alumnos que ven un video, se involucraron en un 74% en realizar el cuestionario que está asociado (Kovacs, G. (2016).

La preparación de un video interactivo por parte del docente incrementa la participación del alumno, quien muchas veces está acostumbrado a recurrir a la web y las redes sociales para obtener información e interacción. Según la observación realizada por diversos estudios cualquier estímulo tiende a desaparecer después de unos 10 minutos, y como resultado los alumnos demandan nuevos contenidos después de ese corto espacio de tiempo. Uno de los beneficios de este tipo de material de no muy larga duración es que permite romper la instrucción directa y demasiado prolongada en secuencias más cortas, y más digeribles (Amo, D, Santiago, R., 2017).

Es de vital importancia poseer plataformas donde distribuir este tipo de material audiovisual, no solo para su visualización, sino también para la recogida de datos sobre su uso. Se debe articular pues dónde montar estos materiales. Surge por ello de esta investigación, indagar sobre diversos Entornos Virtuales Flexibles (EVF), pues estos 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-Peñalvo, Conde-González, 2015). Además, tener en cuenta que, mostrar la información en formato de video no conducirá automáticamente a un aprendizaje en profundidad (Karppinen, P., 2005) y que el esfuerzo del docente debe estar en el diseño pedagógico bien fundamentado cuando desarrolla videos para que realmente sea una herramienta efectiva en contextos educativos.

Es muy importante alinear los objetivos de aprendizaje a la evaluación, pues la alineación constructiva implica el diseño de aprendizaje y evaluación entorno al aprendizaje destinado a resultados u objetivos. Si se logra que las actividades de aprendizaje que se han previsto se articulen en forma explícita con los resultados de aprendizaje, nuestros alumnos podrán construir su propio aprendizaje (Biggs, J., Tang, C., 2007).

El uso de videos enriquecidos con preguntas se lo suele denominar videoquiz, pues se usa el medio audiovisual y se le asigna una serie de preguntas que pueden ser de tipo múltiple-choice, verdadero/falso o preguntas abiertas de comprensión. Algunas herramientas de este tipo que se han analizado son Edpuzzle, Playposity, TedEx son de gran ayuda pues nos proporcionan datos muy interesantes sobre el nivel de conocimiento de cada uno de nuestros alumnos, cuando se siguen ciertas pautas para su uso, permitiendo estructurar el conocimiento y guiar hacia un aprendizaje eficaz y que como docentes nos permita armar actividades de orden superior en función de la Taxonomía propuesta por Bloom: analizar que saben antes de comenzar el contenido, evaluar sus conocimientos durante y después de las actividades y crear materiales y productos de aprendizaje fruto del proceso que hayan llevado a cabo (Amo, D, Santiago, R., 2017).

Muchas veces es probable que las respuestas a la hora de completar actividades propuestas por el docente no sea la que esperamos, es por ello que se puede recurrir a la gamificación también usando videoquiz, para ello se debe tener en cuenta una serie de pautas relacionadas a técnicas de diseño de juego principalmente de recompensas en entornos virtuales. Todo ello permitirá estimular la participación de los alumnos, hacer más entretenidas las actividades difíciles y aburridas, promover la perseverancia y la participación, generar compromiso en la tarea y permitir que el alumno pueda relacionar sus conocimientos nuevos con los que ya posee y re-ajustarlos con la nueva información adquirida y así alcanzar los objetivos pedagógicos (Amo, D, Santiago, R., 2017).

Todos estos desempeños pueden ser visualizados mediante informes que ayudan a determinar el perfil del alumno, paneles de visualización de tendencias, o gráficos que permitan dar valor a las notas obtenidas y que quedarán representadas mediante estos paneles, también se puede incorporar el uso de mapa de calor, que puede ser de gran utilidad para el docente para los listados con notas tradicionales. El mapa de calor consiste en mostrar mediante una gama de colores las distintas calificaciones obtenidas por los alumnos, dando un significado a cada color según su respuesta: correcta/incorrecta, la demora en dar la respuesta, la completitud o no de la consigna, etc. Se ha estudiado algunas de estas herramientas donde el mapa de calor está implementado y se ha experimentado en el aula tal es el caso de las aplicaciones: Kahoot, Socrative y Quizizz que utilizan la gamificación mediante el uso de dispositivos móviles, para este caso concreto la experimentación se hizo mediante el uso de celulares en tiempo real en clase y fuera de ella como tareas para el hogar (homework). De esta experiencia es posible hacer un uso similar también para los videoquiz.

Toda la información obtenida de esta recogida de datos, se los puede obtener en una simple hoja de

cálculo o también mediante informes de los entornos virtuales como LMS (Learning Management System) que combinados con LRS (Learning Record System) permiten obtener un seguimiento del e-learning aplicando estándares tales como SCORM y el nuevo estándar Experience Api o xApi, que ha evolucionado respecto del anterior pues no estaba preparado para el aprendizaje informal, social, ubicuo y móvil y que es parte de nuestra investigación actual. Las analíticas del aprendizaje emergieron inicialmente para estudiar datos a posteriori, describiendo resultados y aportando diagnósticos, pero su evolución hoy en día ayuda a identificar qué alumnos avanzan con dificultad, las áreas más débiles, identificar los factores que afectan al éxito de los alumnos y sus necesidades individuales. Además, los modelos de datos nos dan la posibilidad de identificar tendencias y patrones que ayuden a los docentes a saber diseñar, cómo apoyar y asistir en forma más personalizada para permitir intervenir en casos de dificultad en el aprendizaje, evitar abandonos innecesarios desmotivación y la posibilidad de repetición de evaluaciones y lograr un aprendizaje más reflexivo.

2. Líneas de investigación y desarrollo

La línea I/D que presenta este artículo, consiste en la aplicación de las analíticas del aprendizaje en el uso de videos educativos diseñados para dispositivos móviles; esto, con el fin de recoger y almacenar datos relacionados con la experiencia de aprendizaje de los alumnos mediante el uso de estos videos, permitiendo identificar tendencias y patrones de comportamiento, como así también obtener algunos indicadores de rendimiento. Consideramos de gran relevancia analizar las huellas que dejan nuestros alumnos al usar estos videos enriquecidos con cuestionarios, pues permite recabar información para ayudar a la toma de decisiones de los docentes durante el proceso de enseñanza-aprendizaje, permitiendo analizar su comportamiento para así acercarnos a su nivel de conocimiento. De esta manera se les estará guiando hacia un aprendizaje eficaz y significativo, se podrán mejorar actividades o lecciones y en función de las analíticas obtenidas, se podrán diseñar actividades de diferenciación dentro del grupo, descartar recursos poco vistos o re formular cuestionarios poco claros.

3. Resultados Esperados

Objetivos ya alcanzados:

- Identificar las competencias específicas para el aprendizaje de la Programación.
- Definir estrategias en el uso de un Entorno Virtual Flexible para el desarrollo de competencias para el aprendizaje de la Programación.
- 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.

Parte del equipo de I/D ha estudiado diversas herramientas y ha llevado a cabo distintas experiencias en la asignatura Programación Procedural, que corresponde al primer año de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información de la Facultad de Ciencias Exactas de la Universidad Nacional de San Juan (Llarena, Díaz, 2018). (Vera C., Llarena M., Díaz M., 2017), como así también en la asignatura Sistemas de Información II, correspondiente al 4to año de Licenciatura en Sistemas de Información de la Facultad de Ciencias Exactas de la Universidad Nacional de San Juan. Siendo las experiencias muy satisfactorias para seguir con el curso de esta línea de I/D.

Algunos de los objetivos a concretar sobre la línea de investigación específica para este artículo son:

- Analizar los estándares avanzados de aprendizaje distribuido SCORM y xAPI.
- Proponer recomendaciones para diseñar herramientas para dispositivos móviles útiles para la analítica del aprendizaje.
- Experimentar qué modelo pedagógico se adapta mejor al uso de las herramientas investigadas.
- Evaluar el trabajo de campo realizado.
- Sistematizar el uso de las analíticas de aprendizaje para que el docente pueda

evaluar el proceso de aprendizaje de sus alumnos.

4. Formación de recursos humanos

Cabe destacar que en este proyecto de I/D existen tres tesis de Posgrado, donde una de ellas consiste en establecer estrategias para diseñar recursos didácticos para dispositivos móviles, y evaluar su uso e impacto. Dicha tesis está en etapa de escritura para obtener el grado de magíster de la Maestría en Informática dictada en la Universidad Nacional de San Juan.

5. Bibliografía

Amo, D., & Santiago, R. (2017). Learning analytics: la narración del aprendizaje a través de los datos. UOC.

Biggs, J., & Tang, C. (2007). Teaching for quality learning at university Maidenhead. Berkshire, UK: McGraw-Hill Education.

Bravo Ramos, J. L. (2000). El video educativo. Madrid: ICE de la Universidad Politécnica de Madrid, 3.

Humanante-Ramos, P. R., García-Peñalvo, F. J., & Conde-González, M. Á. (2015). Personal learning environments and online classrooms: An experience with university students. IEEE Revista Iberoamericana de tecnologías del aprendizaje, 10(1), 26-32.

Karppinen, P. (2005). Meaningful learning with digital and online videos: Theoretical perspectives. AACE journal, 13(3), 233-250.

Kolb, D. A. (1984). The process of experiential learning. Experiential learning: Experience as the source of learning and development, 20-38.

Kovacs, G. (2016, April). Effects of in-video quizzes on MOOC lecture viewing. In Proceedings of the third (2016) ACM conference on Learning@ Scale (pp. 31-40). ACM.

Llarena, M. G., & Díaz, M. (2018). Los entornos virtuales flexibles para el desarrollo de competencias en el área de algoritmos y lenguajes de programación. In XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).

Vera, C., Llarena, M. G., & Díaz, M. (2017). Entornos personales como apoyo para el aprendizaje y desarrollo de competencias: experiencia con alumnos de informática. In XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET, La Matanza 2017).

Zankov, L. V. (1977). Teaching and development: a soviet investigation. ME Sharpe.

Enfoques metodológicos para la enseñanza de la programación en la escuela secundaria

Federico Amigone[1] Ruben Ramirez[2] Daniel José Dolz[1]
Gerardo Parra[1] Jorge Rodríguez[1]

fede.amigone@fi.uncoma.edu.ar, rhubenariel@gmail.com, ddolz@fi.uncoma.edu.ar, gparra@fi.uncoma.edu.ar, j.rodrig@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 la informática en la escuela secundaria está transitando una profunda transformación. Existe un amplio consenso en incorporar contenidos sobre Ciencias de la Computación en las propuestas curriculares para la educación obligatoria.

El área de conocimiento Algoritmos y Programación ocupa un lugar importante en las iniciativas y propuestas curriculares que se elaboran en este contexto. En los últimos años se han desarrollado entornos dedicados específicamente a la enseñanza de la programación. La definición de enfoques para la enseñanza y aprendizaje de la programación en el ámbito escolar es una pieza clave para el desarrollo y consolidación de los procesos que buscan aproximar las Ciencias de la Computación a la Escuela Secundaria.

En este trabajo se presenta una Línea de Investigación que busca producir conocimiento para favorecer los procesos de enseñanza y aprendizaje de la programación en ambientes escolares. Esta línea se basa en investigaciones recientes acerca de la enseñanza de la programación, en particular en el campo de Embodied Programming, enfoques que proponen la enseñanza de la programación a partir de la modificación de código existente y estudios sobre cómo facilitar los procesos de transición de lenguajes basados en bloque a lenguajes basados en texto.

Palabras Clave: Ciencias de la Computación, Enseñanza de la Programación.

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

Aún con diferente grado de avance en los distintos países, existe un amplio consenso

en incorporar contenidos de Ciencias de la Computación en las propuestas curriculares para la educación obligatoria . En este sentido, la enseñanza de la informática en la escuela secundaria está transitando una profunda transformación[3, 6].

En los últimos años se han producido numerosos aportes y se han desarrollado variados entornos dedicados específicamente a la enseñanza de la programación. Es por ello que el área de conocimiento Algoritmos y Programación ocupa un lugar muy importante en las iniciativas y propuestas curriculares que se elaboran en este contexto. La definición de enfoques para la enseñanza y aprendizaje de la programación en el ámbito escolar es una pieza clave para el desarrollo y consolidación de los procesos que buscan aproximar las Ciencias de la Computación a la Escuela Secundaria.

En este trabajo se presenta una Línea de Investigación que tiene como objetivo producir conocimientos y definir enfoques tendientes a favorecer los procesos de enseñanza y aprendizaje del área Algoritmos y Programación en ambientes escolares.

Esta línea se fundamenta y construye sobre investigaciones recientes en el campo de la enseñanza de la programación. En particular, se consideran las perspectivas construidas en el campo de Embodied Programming, los enfoques que proponen la enseñanza de la programación a partir de la modificación de código existente y estudios acerca de cómo facilitar los procesos de transición de lenguajes basados en bloque a basados en texto.

2. Línea de Investigación

En esta sección se presentan los ejes principales que componen la Línea de Investigación.

2.1. Enfoque didáctico disciplinar

La definición de un enfoque requiere algunas definiciones previas, entre ellas, el diseño del andamiaje didáctico sobre el que se articularán las metodologías concretas.

El modelo propuesto parte de considerar al estudiante como un sujeto vinculado a un contexto que se encuentra atravesado por la aplicación instrumental de los campos disciplinares que se estudian. De esta forma, se persigue replantear el sentido común imperante y percibido del mundo ofreciendo herramientas cognitivas que maximicen la capacidad de análisis de la realidad que lo rodea. Se intenta que el alumno domine el campo disciplinar para que dicho dominio lo habilite a una mejor comprensión de los procesos tecnológicos que urden el funcionamiento de los espacios sociales, laborales y económicos de su contexto.

2.1.1. Punto de partida

El enfoque triangula la subjetividad, el contexto y el campo disciplinar. Desde esa terna se desprenden instanciaciones metodológicas calibradas y funcionales. Existe mucha información que se infiere a partir de la relación ternaria propuesta. En principio, pone al sujeto en una relación contextual con su medio y permite determinar el grado de conocimiento y fricción con las manifestaciones tecnológicas que serán estudiadas.

2.1.2. Recorrido constructivista

Independientemente del grado de conocimiento disciplinar del sujeto, casi siempre podrán establecerse relaciones entre éste y los efectos de la tecnología en su contexto. Esas relaciones de problematización son utilizadas como disparadores y puntos de partidas para el recorrido didáctico. Desde la triangulación como punto de partida, comienza un recorrido que parte de un elevado nivel de abstracción técnica y a través de transposiciones didácticas, se comienza a aproximar acercamientos disciplinares concretos. Se llega al final del recorrido, cuando el sujeto consigue construir un artefacto computacional que sea vinculante a su contexto específico.

2.1.3. Dispositivos didácticos

Son diseñados en función de los niveles de cohesión entre las relaciones sujeto, contexto y disciplina. Son dispositivos pensados a partir de la problematización propuesta, y cuya aplicación permiten desarrollar, con una propuesta constructivista, determinadas nociones del campo de estudio. Estos dispositivos se aplican con un sentido incremental en su acceso concreto a la disciplina.

2.2. Ejes de trabajo

2.2.1. Embodied Programming

Embodied Learning constituye un paradigma de aprendizaje que enfatiza el uso del cuerpo en la práctica educativa, se basa en la premisa de que existe un vínculo inseparable entre cuerpo y mente en el proceso de aprendizaje [5, 2].

Desde este paradigma se supone que el aprendizaje ocurre como fenómeno interactivo entre sujetos, sus cuerpos, los artefactos en el entorno físico y las acciones que realizan con esos artefactos. Este enfoque constituye un recurso innovador para aprender conceptos abstractos como los relacionados al desarrollo de algoritmos[8, 2].

Surrogate Embodiment se refiere a un tipo de interacción donde los movimientos realizados por una persona sobre un escenario son dirigidos explícitamente por los estudiantes. Investigaciones recientes sostienen que existe cierta similitud entre Surrogate Embodiment y las prácticas computacionales que se ponen en juego durante el desarrollo de algoritmos. En el sentido de que la manipulación explícita de un sustituto se asemeja a los procesos de pensamiento de un programador [8].

En el contexto de esta Línea de Investigación se reconoce la importancia de definir un modelo que considere diferentes formas de representación posibilitando la construcción gradual de la capacidad de elaborar abstracciones y manipular conceptos y prácticas fundamentales del área Algoritmos y Programación. Se parte de experiencias concretas que permitan, más adelante, construir conocimientos abstractos.

Surrogate Embodiment con representación física: las acciones de un substituto externo, preferiblemente el docente, son dirigidas explícitamente por los estudiantes. La actividad se desarrolla sobre un escenario montado físicamente.

Surrogate Embodiment con representación gráfica: La actividad se desarrolla sobre un escenario representado en forma gráfica, los estudiantes indican como mover un objeto por el escenario.

Surrogate Embodiment con representación imaginaria: la manipulación del sustituto se realiza sobre un escenario imaginario, un modelo mental.

Direct Embodiment con representación imaginaria: las acciones se realizan sobre un modelo imaginario del escenario sin participación de un sustituto.

2.2.2. Enseñanza de la programación a partir de la modificación de código existente

Estudios recientes acerca de la enseñanza de conceptos y prácticas sobre algoritmos y programación sostienen partir de la modificación de código existente constituye un enfoque adecuado para estudiantes sin formación previa en el área de conocimiento[1, 7].

Pair Programing es una práctica originada en la industria del software como parte de las metodologías de Extreme Programming. Es una práctica donde dos programadores, driver y navigator, trabajan colaborativamente en diseño de algoritmos, codificación y prueba de programas.

Pair Programming demuestra ser un método efectivo para contribuir al mejoramiento de la calidad de los procesos de enseñanza y aprendizaje de la programación[9].

En esta Línea de Investigación se propone desarrollar un modelo didáctico que potencie los procesos reelaboración de código a partir de la aplicación de técnicas de Pair Programming.

Se plantea un modelo en espiral de complejidad creciente donde en cada iteración se presenta la necesidad de adecuar una porción de código a una funcionalidad diferente.

En cada cada ciclo, los estudiantes trabajando en parejas, analizan un programa existente, identifican porciones de código que deben ser modificadas e introducen cambios. Luego de algunas iteraciones, cuando el aprendizaje está consolidado, la pareja está en condiciones de enfrentar desafíos más complejos.

En el modelo propuesto se puede utilizar para estructurar cada iteración la siguiente secuencia de actividades:

- Conocer y comprender: Se presenta una pieza de código, las parejas estudian el programa para dar cuenta sobre que problema resuelve y cómo lo hace. Esta actividad, además, se constituye en momento de activación del par de programación.
- Práctica guiada: Se realiza con la colaboración del docente quien guía una primer intervención sobre el código.
- Colección de desafíos: Se propone al par de programación modificar código existente para resolver una colección de desafíos, dónde la complejidad se incrementa al avanzar por el espiral.
- Evaluación formativa: Se despliegan actividades tendientes a mejorar la compresión acerca del logro de los estudiantes respecto las metas.
- Formalización conceptual: Los aprendizajes logrados se expresan formalmente e integra lo aprendido a un marco conceptual.
- Evaluación sumativa: Se realiza al finalizar las iteraciones con intención de comprobar los resultados obtenidos.

2.2.3. Transición de programación basada en bloques a basada en texto

El enfoque de programación por bloques demuestra ser una opción efectiva para dar los primeros pasos.

En el transcurso de la escuela secundaria los estudiantes deberán programar usando lenguajes basados en texto. El paso a un entorno de programación tradicional, como Netbeans+java, puede resultar frustrante [4].

En esta Línea se propone estudiar entornos que favorezcan y faciliten la transición de programación basada en bloques a programación basada en texto.

En tal sentido se han realizado experiencias satisfactorias con herramientas del tipo mixtas con ambientes tipo de caja de arena.

Por herramientas de tipo mixtas se hace referencia a las que posibilitan trabajar con bloques visualizando en lo inmediato el código producido automáticamente, y viceversa. Se trabaja en la especificación de alguna herramienta que permita este paso. Idealmente dicha herramienta debería contener:

- Sandbox: Proponemos que el ámbito de aplicación de la herramienta no sea general, sino específico contenido en un intérprete de preferencia visual, dentro de la misma herramientas.
- Multilenguaje: la herramienta debería proveer la posibilidad de generar e interactuar con más de un lenguaje de programación.
- Bidireccional: debería ser posible generar código a partir de bloques, y viceversa.
- Autoevaluativa: no debería ser necesaria la ayuda de un docente para que el estudiante sepa si su programación resolvió el problema o no; aunque siempre será necesaria una instancia de evaluación de eficiencia, elegancia y posibles alternativas de solución diferentes.

- Instalable: esta línea siempre tiene en cuenta estudiantes sin acceso permanente a internet, con lo que una herramienta puramente web-enabled no estaría al alcance de toda la comunidad educativa.

Se considera importante avanzar en la obtención de una herramienta de estas características, tanto desde la especificación como desde la implementación.

3. Resultados

En el contexto de esta Línea de Investigación se espera obtener los siguientes resultados:

- Definir un modelo didáctico que articule las diferentes tipos de Embodied Programming.
- Proponer un modelo de la enseñanza basado en modificación de código existente en espiral de complejidad creciente.
- Construir especificaciones para herramientas que faciliten la transición de programación basada bloques a texto.
- Avanzar en la articulación de un enfoque general que integre la gama de perspectivas presentadas en este trabajo.

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] K. Brennan and M. Resnick. New frameworks for studying and assessing the development of computational thinking. In *Proceedings of the 2012 annual meeting of the American Educational Research Association, Vancouver, Canada*, 2012.

[2] C. L. Fadjo. *Developing computational thinking through grounded embodied cognition.* PhD thesis, Teachers College, 2012.

[3] K-12 Computer Science Framework Steering Committee. *The K–12 Computer Science Framework.* ACM, 2016.

[4] M. Kölling, N. C. Brown, and A. Altadmri. Frame-based editing: Easing the transition from blocks to text-based programming. In *Proceedings of the Workshop in Primary and Secondary Computing Education*, pages 29–38. ACM, 2015.

[5] P. Kosmas and P. Zaphiris. Embodied cognition and its implications in education: An overview of recent literature. *International Journal of Educational and Pedagogical Sciences*, 2018.

[6] F. Sadosky. *CC – 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas* . Fundación Sadosky, Argentina, 2013.

[7] S. Sentance, J. Waite, and M. Kallia. Teachers' experiences of using primm to teach programming in school. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, pages 476–482. ACM, 2019.

[8] W. Sung, J. Ahn, and J. B. Black. Introducing computational thinking to young learners: practicing computational perspectives through embodiment in mathematics education. *Technology, Knowledge and Learning*, 22(3):443–463, 2017.

[9] D. Tsompanoudi, M. Satratzemi, and S. Xinogalos. Distributed pair programming using collaboration scripts: An educational system and initial results. *Informatics in Education*, 14(2):291–314, 2015.

EscuelasTIC. El pensamiento computacional en la escuela

Claudia Queiruga, Claudia Banchoff Tzancoff, Paula Venosa, Sofía Martin, Vanessa Aybar Rosales, Soledad Gomez, Isabel Kimura
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
{claudiaq, cbanchoff, pvenosa, vaybar, sgomez}@info.unlp.edu.ar, {smartin, ikimura}@linti.unlp.edu.ar

Resumen

En los últimos años las experiencias de incorporación de tecnologías digitales en las aulas de las escuelas de nuestro país se ha desarrollado aceleradamente, sin embargo, en su gran mayoría el enfoque adoptado es el de "informática de usuario" que hace hincapié en el conocimiento y la capacidad de utilizar las computadoras, y la tecnología relacionada con ellas, de manera eficiente con un enfoque "utilitario" en algunos casos e "integrador" en otros, sin abordar las especificidades de la informática como disciplina.. Si bien ha habido un gran avance, la sociedad actual, caracterizada por estar cada vez más conectada, en donde la mayoría de las actividades cotidianas están mediadas por objetos tecnológicos requiere conocer, comprender e intervenir este nuevo mundo digital. El enfoque del "pensamiento computacional", aunque aún no está plenamente consolidado como nuevo paradigma de enseñanza de la informática en la escuela, amplía las miradas instrumentales sobre los saberes de la informática en el sistema educativo obligatorio. Este artículo presenta algunos avances de la línea de investigación, denominada, "EscuelasTIC", que se viene llevando a cabo desde hace varios años, dando continuidad a la línea de trabajo "PROGRAMAR en la Escuela".

Palabras clave: Informática, enseñanza de programación, robótica educativa, pensamiento computacional, formación docente, TIC.

Contexto

La línea de investigación "EscuelasTIC" 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). Esta línea de trabajo, presentada en eventos anteriores (Queiruga C., et al., 2018), articula los procesos de investigación con actividades de extensión que promueven la creación con tecnologías digitales en el aula, con especial atención en la incorporación del enfoque del pensamiento computacional como paradigma de enseñanza de la Informática en la escuela. La formación docente ocupa un espacio central y este equipo de investigadoras trabajó en el diseño curricular de la carrera de "Especialización docente en didáctica de las Ciencias de la Computación" de la provincia de Buenos Aires, que comenzó a dictarse en septiembre de 2018.

Introducción

La sociedad actual, denominada por algunos autores Sociedad de la Información o Sociedad del Conocimiento y, por Manuel Castells[1] "Capitalismo Informacional" describe una sociedad mediada por objetos digitales, Esta nueva sociedad, requiere entre otras cosas, entender cómo funcionan estos nuevos objetos y cómo adoptarlos críticamente. En el campo educativo se plantea la formación de ciudadanas y ciudadanos 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 (Queiruga C. et al, 2018).

El uso de TIC se ha ido incorporando ampliamente a las prácticas educativas en los distintos niveles de la escolaridad obligatoria de nuestro país, primando el enfoque de "informática de usuario" que hace hincapié en el conocimiento y la capacidad de utilizar las computadoras, y la tecnología relacionada con ellas, de manera eficiente con un enfoque utilitario en algunos casos e "integrador" en otros (Levis M., 2007). En los últimos años se ha ido modificando este enfoque y en los diseños curriculares del sistema obligatorio de algunas provincias, comienza a aparecer el enfoque del pensamiento computacional como nuevo paradigma de enseñanza de la informática. Jeannette Wing, fue quien en 2006 acuñó el término "pensamiento computacional" como abreviatura de "pensar como informáticos", y lo describe como una forma de pensar que promueve habilidades y competencias cognitivas que favorecen maneras de pensar diferentes a la de otras ciencias, con características propias, entre ellas la descomposición de problemas en subproblemas, la abstracción de casos particulares, los procesos de diseño, implementación y prueba de lógicas algorítmicas, entre otras (Wing J., 2006). A modo de ejemplo, la provincia de Buenos Aires incorporó en el año 2018, en el diseño curricular de educación primaria (DGCyE/Diseño curricular para la educación primaria de la provincia de Buenos Aires, 2018) las ideas del pensamiento computacional y la intención de incorporarlas transversalmente en las diferentes áreas curriculares.

En el caso de la educación de nivel secundario en la provincia de Buenos Aires, la asignatura "NTIC" (DGCyE/Diseño curricular para la educación secundaria ciclo superior, 2010), ubicada en 4to. año, sigue siendo el único espacio curricular donde se abordan algunos conceptos relacionados a las Ciencias Informáticas. Esto muestra a la informática como un espacio de conocimiento alternativo, aún no legitimado en el ámbito de la educación escolar, con escasa relevancia en la formación de las y los estudiantes.

Incorporar en los diseños escolares la disciplina Informática desde el enfoque del pensamiento computacional constituye un logro y un espacio de seguridad para el desarrollo de procesos cognitivos vinculados al razonamiento lógico, el cual permite: predecir, analizar y explicar problemas; formular algoritmos; descomponer problemas en partes más simples; abstraer para manejar la complejidad; generalizar mediante el descubrimiento de patrones y similitudes y evaluar soluciones. Estas ideas son ampliamente aplicadas en la resolución de problemas usando computadoras y en la comprensión de los sistemas más allá del espacio escolar.

---

1 Capitalismo informacional, término introducido por Manuel Castells para identificar el actual período del capitalismo que hace referencia al conocimiento digital como parte de todos los procesos productivos.

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 más recientemente el "Plan Nacional Integral de Educación Digital" (PLANIED, sf), el proyecto "Secundaria 2030" (Secundaria 2030, sf), que han puesto el acento en ampliar la formación tecnológica, educando a los jóvenes en la comprensión de cómo funcionan las tecnologías digitales y sus principios fundantes, de manera de convertirlos en agentes creativos del mundo digital y no consumidores pasivos. En este sentido el pensamiento computacional y los conceptos de Informática han comenzado a incorporarse en los nuevos currículos escolares reemplazando las ideas vinculadas al uso eficiente.

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 en Informática 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 y, a la construcción de contenidos con tecnologías digitales pertinentes al espacio escolar. Se está trabajando con docentes y estudiantes de 15 escuelas secundarias y 3 escuelas primarias del distrito escolar de La Plata, Berisso y Ensenada.

Respecto a la formación docente, en el año 2017 se desarrolló el diseño curricular de la "Especialización docente en didáctica de las Ciencias de la Computación" de la provincia de Buenos Aires (Queiruga C., et al., 2018). Esta carrera, que comenzó a dictarse en septiembre de 2018 en tres sedes de la provincia de Buenos Aires, constituye un resultado de la investigación en el campo y de las experiencias en el territorio. Los espacios promovidos por los proyectos de extensión son de suma relevancia para esta línea de investigación, dado que en los mismos se realizan actividades con los docentes en los que se trabajan temas y materiales que resultan en insumos para el desarrollo de sus clases. Se trata de espacios de formación específicos para docentes de los distintos niveles sobre las nuevas didácticas en relación a la Informática (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 niñas, niños 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 la 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, niñas 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 también parte de las actividades en esta línea de investigación y sus distintos ejes íntimamente vinculados a la formación en ciudadanía digital.

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 el enfoque del pensamiento computacional y el uso seguro y responsable de las tecnologías digitales; atendiendo a la formación de las y los ciudadanos del siglo XXI que requieren de nuevas habilidades y destrezas vinculadas a diseñar, crear e innovar con los nuevos 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 favoreciendo trabajar la Informática en la escuela en todos los niveles aplicando conceptos de gammification[2] 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 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. 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 en escuelas; otras están en procesos de prueba y evaluación. Asimismo se trabaja articuladamente con algunas asignaturas de las carreras de la Facultad de Informática en el desarrollo de distintas herramientas.

En las actividades de esta línea de investigación, este equipo de investigadoras ha participado en 3 proyectos impulsados por la Fundación Sadosky cuyo objetivo es aportar desde una visión federal a la incorporación de la disciplina Informática en las escuelas: 1) 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 2019. 2) El diseño curricular y dictado de la primera "Especialización docente en didáctica de las Ciencias de la Computación" junto con el ISFD Nº 95 de La Plata, destinada a docentes de nivel medio de la provincia de Buenos Aires. cuya formación de base es matemática, física, química y tecnología. 3) La implementación del diseño curricular de la "Especialización docente en didáctica de las Ciencias de la Computación" dictada en el ISFD Nº 95 de La Plata. Esta carrera cuenta con una matricula de

2 El término *Gamification* hace referencia al uso del diseño y la mecánica de juegos para mejorar los contextos no relacionados con los juegos.

120 docentes. Estos proyectos, nos proveen evidencias científicas con base empírica para continuar indagando sobre las formas de incorporar esta disciplina en los sistemas de educación formales y obligatorios.

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 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.

Referencias

Díaz J., Banchoff Tzancoff C., Queiruga C. y Martin S. (2014). Experiencias de la Facultad de Informática en la Enseñanza de Programación en Escuelas con Software Libre. 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.

Dirección General de Cultura y Educación de la Provincia de Buenos Aires . Diseño curricular para la educación primaria: primer ciclo y segundo ciclo; coordinación general de Sergio Siciliano. - 1a ed. - La Plata (2018). ISBN 978-987-676-095-9.

Dirección General de Cultura y Educación de la provincia de Buenos Aires / Diseño Curricular para la Educación Secundaria Ciclo Superior. ES4: Nuevas Tecnologías de la Información y la Conectividad / coordinado por Claudia Bracchi. -1a ed.- La Plata (2010). ISBN 978-987-1266-98-2 .

Levis D. (2007) “Enseñar y aprender con informática/ enseñar y aprender informática. Medios informáticos en la escuela argentina” en Cabello, R. y Levis D., edits. Tecnologías informáticas en la educación a principios del siglo XXI Buenos Aires: Prometeo.

PLANIED: Plan Integral de Educación Digital del Ministerio de Educación y Deportes de la Nación Argentina (s.f).

Program.AR: programa de la Fundación Sadosky, Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación Argentina (s.f).

Queiruga C., Banchoff Tzancoff C., Venosa P, Gómez S., Morandi G. (2018). Ciencias de la Computación y escuelas ¿una didáctica específica?. En CACIC 2018 (Congreso Argentino de Ciencias de la Computación). UNICEN, Tandil, Argentina. Octubre de 2018. Libro de Actas XXIV CACIC 2018. pp 1040-1049. ISBN 978-950-658-472-6.

Queiruga C., Banchoff Tzancoff C., Venosa P., Martin S., Aybar Rosales V., Gomez S., Kimura I. (2018). EscuelasTIC: las tecnologías digitales en las aulas. UNNE, Corrientes, Argentina. Abril de 2018. Libro de Actas XX WICC 2018. . pp 475-489. ISBN 978-987-3619-27-4.

Queiruga C., Banchoff Tzancoff C., Martin S., Aybar Rosales V., López F., Kimura I. y Gómez S. (2017). PROGRAMAR en la Escuela: Nuevos Desafíos en las Aulas. XIX 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).

Secundaria 2030: Transformar la Secundaria para transformar vidas (s. f).

Wing, J. M. (2008). Computational thinking and thinking about computing: Philosophical Transactions of The Royal Society A, vol. 366, 3717–3725.

Wing, J. M. (2006). Computational Thinking. Communications of the ACM, vol. 49, 33-35.

Estrategia de Divulgación y Enseñanza para Fomentar el Interés en las Tecnologías Espaciales

Andres Torti, Leonardo Anchino, Leandro Chiappero, Marco Miretti, Emmanuel Dovis, Emanuel Bernardi, Juan Calloni, Jorge Tomé, Mario Alberto Berón, Rodolfo Podadera†
Grupo de Investigación en Robótica y Control (GIROC)
Universidad Tecnológica Nacional
Facultad Regional San Francisco (UTN-FRSFCO)
San Francisco, Córdoba, Argentina
†rodolfopodadera@gmail.com

RESUMEN

Mediante el presente proyecto, se propone analizar el impacto y el efecto multiplicador que proporciona la incorporación de una plataforma tecnológica de vanguardia, como son los dispositivos CANSAT, en el sistema educativo de nivel medio. Dicho de otro modo, la inclusión de las plataformas CANSAT manifiesta el deseo de posicionar a las instituciones, universitaria y de nivel medio, en el punto más avanzado de la tecnología educativa, con el objetivo de despertar e incentivar vocaciones científicas y tecnológicas entre los jóvenes alumnos.
Se entiende como plataforma CANSAT a un pequeño satélite[1], no orbital, con las dimensiones de una lata de gaseosa, como medio diseñado especialmente para permitir a los alumnos de nivel medio insertarse en el mundo de las tecnologías espaciales a través de la programación, la electrónica, el uso de estadísticas, la interpretación de datos, etc.

Palabras clave: estudio, educación técnica, cansat, satélites, ingeniería.

CONTEXTO

Esta propuesta de investigación se enmarca dentro de las áreas Electrónica Digital, Informática, Sistemas Embebidos, Control de Sistemas y Comunicaciones del Departamento de Ingeniería Electrónica de la Universidad Tecnológica Nacional, Facultad Regional San Francisco y del establecimiento educativo de nivel medio IPET N° 50 "Emilio F. Olmos". Además, el proyecto ha sido homologado por el Instituto Nacional de Educación Tecnológica (INET), cuyo expediente es EX2018041815675-APN-INET#ME, y su título "Proyecto CANSAT EFO-UTN". Por otro lado, las actividades involucradas en este se llevarán a cabo por integrantes de ambas instituciones. Específicamente, por directivos y docentes del establecimiento de nivel medio, y por los integrantes del grupo de Investigación en Robótica y Control (GIROC), compuesto de becarios alumnos, docentes investigadores y becarios doctorales.

1. INTRODUCCIÓN

El proyecto consta de los siguientes aspectos: TÉCNICOS: Diseño, construcción, lanzamiento y recuperación. Diseño y construcción de los módulos de control, medición, transmisión y recepción. Lanzamiento del sistema mediante diversas técnicas (globo cautivo, cohete, dron), recuperación del dispositivo y análisis e interpretación de los datos recolectados. PEDAGÓGICOS: determinación de las áreas curriculares y actividades formativas para el diseño y construcción. Determinación de las áreas curriculares y actividades formativas para el lanzamiento, interpretación de los datos

1 Si bien se los denomina «satélites», no lo son en el sentido estricto de su definición como cuerpo que gira alrededor de un planeta.

y recuperación del sistema. VINCULARES: entre la UTN, Facultad Regional San Francisco y el IPET N° 50, como experiencia de articulación. Construcción de lazos con instituciones establecidas a nivel nacional e internacional. Siendo estas el Instituto Civil de Tecnología Espacial (ICTE) [4] y el Consorcio de Universidades de Ingeniería espacial UNISEC-global [3]. Promoción y transferencia de la experiencia en otras instituciones. Dichas acciones comprenden la visita a establecimientos educativos, la publicación gratuita de los diseños electrónicos realizados y la difusión en los medios de comunicación. Se prevé que el proyecto tenga dos niveles de participación a saber: Nivel Inicial: la estructura del CANSAT se asienta en su mayoría sobre una plataforma CANSAT [5-9] "lista para usar", desarrollada por alumnos de la UTN, Facultad Regional San Francisco. De esta manera se facilita el ingreso y participación de las escuelas no técnicas. Nivel Avanzado: diseño y construcción sobre una plataforma base CANSAT, a cargo de la escuela secundaria. Para ello se dispondrá de una serie de diseños gratuitos y módulos elementales sugeridos desde el Departamento de Ingeniería Electrónica de la UTN, Facultad San Francisco. Ambos niveles podrán utilizar distintos tipos de lanzamiento (globo cautivo, cohete, dron), con diseño propio o comercial. Asimismo, se prevé tomar como elemento base al reglamento de la competencia CanSat 2019 del CanSat Leaders Training Program (CLTP), [1, 2], el cual establece los modos de participación, altura de lanzamiento, pesos y mediciones

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 necesarias para el desarrollo del sistema y realización del IDE (Integrated Development Environment, o su traducción al español ambiente de desarrollo integrado).
- Estudio de la disponibilidad y oferta local de los dispositivos a integrar en el sistema.
- Evaluación de la estrategia de procesamiento de los datos captados por el sistema de sensores.
- Desarrollo del software y hardware necesarios para los siguientes sistemas: Adquisición de datos, control y comunicación con el CANSAT.
- Generación de documentación adecuada para la discusión interna de resultados.
- Divulgación de los resultados parciales y finales de la investigación.

3. OBJETIVOS

Objetivo General:

- Fomentar el desarrollo de las capacidades fundamentales de los estudiantes de ambas instituciones. Alentando la vocación científica/tecnológica de los jóvenes, favoreciendo y contribuyendo a la educación espacial y al desarrollo espacial del país. Manteniendo, de esta forma, al cuerpo docente de nivel medio en capacitación continua en relación a nuevas tecnologías.

Objetivos específicos:

- Determinar actividades formativas transversales para ambas instituciones intervinientes:
  - IPET N° 50, que involucren las siguientes asignaturas: matemática, física, educación tecnológica, electrónica digital, informática electrónica, representación gráfica e interpretación de planos, entre otras. Además de aquellas que incluyan temas afines a la programación en alto nivel, la

construcción de hardware, análisis e interpretación de datos.

- UTN Facultad Regional San Francisco, que involucren a las siguientes cátedras: Análisis de Señales y Sistemas, Técnicas Digitales, Medios de Enlace y Sistemas de Comunicaciones, con temáticas atinentes a las curricula de las mismas para así abordar el diseño, construcción, lanzamiento, recuperación, mediciones, análisis, interpretación de datos y divulgación de resultados.

- Cumplir con los objetivos propuestos por las Naciones Unidas en cuanto al desarrollo sostenible.
- Afianzar los lazos de mutua cooperación entre ambas instituciones parte del proyecto.
- Avanzar con el desarrollo de propuestas de alta tecnología en temas educativos para la ciudad y la región.
- Determinar acciones para la promoción del proyecto en otras instituciones educativas y en la sociedad argentina.
- Fortalecer vínculos entre las instituciones educativas participantes y las industrias o empresas productoras de la ciudad y la región
- Divulgar los resultados y capacidades de la investigación.

4. RESULTADOS ESPERADOS

Siendo que el proyecto se encuentra en su fase inicial, durante la ejecución del mismo se pretenden alcanzar los siguientes resultados:

- Incrementar el número de estudiantes en las carreras duras como lo son ingeniería electrónica, sistemas de la información, aeronáutica y espacial.
- Incrementar el número de estudiantes de escuelas técnicas de la región.
- Fortalecer el rendimiento académico en ambas instituciones participantes.
- Posicionar en investigación, desarrollo e innovación a ambas instituciones.
- Disminuir el abandono de estudiantes en ambas instituciones participantes del proyecto.
- Disminuir en la escuela de nivel medio el índice de repitencia de estudiantes.
- Disminuir el tiempo que le insume a un estudiante universitario realizar su carrera.
- Divulgar en el medio donde se trabaja las capacidades del proyecto y sus resultados.
- Formar 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 en Robótica y Control (GIROC), este está conformado por su director Mg. Ing. Rodolfo Podadera, su co-director Dr. Mario Beron, los investigadores formados Ing. Juan Calloni, Ing. Emanuel Bernardi, los investigadores de apoyo Andres Torti, Leonardo Anchino, Emmanuel Dovis, Marco Miretti, el técnico de apoyo Leandro Chiappero, y el Prof. Jorge Tome. Es de destacar, que el ingreso al grupo I+D+i está abierto a nuevos integrantes y desde el mismo se contemplan actividades formativas en relación al tema del proyecto, tanto como temas complementarios. Por último, es importante destacar, que el dispositivo utilizado como base para constituir el presente proyecto forma parte del proyecto final de grado de dos investigadores de apoyo involucrados.

6. BIBLIOGRAFÍA

[1] CLTP, “Cansat Leader Training Program”. URL: http://cltp.info/index.html (visitado 08-03-2019)

[2] European Space Agency, (2019) “CANSAT, 2018-2019 Guidelines”.

[3] UNISEC-Global, “University Space Engineering Consortium”. URL: http://unisec.jp/en.html (visitado 08-032019)

[4] ICTE, “Instituto Civil de Tecnología Espacial”. http://www.icte.com.ar/ (visitado 08-03-2019)

[5] N. Sako, Y. Tsuda, S. Ota, T. Eishima, (2001). “Cansat suborbital launch experiment-university educational space program using can sized pico-satellite”. *acta astronautica*, 48(5-12), 767-776.

[6] S. Bulut, M. Gül, C. Beker, I. İpek, (2013). “Model satellite design for CanSat Competition”. In *2013 6th International Conference on Recent Advances in Space Technologies (RAST)* (pp. 913-917). IEEE.

[7] A. Nylund, J. Antonsen, (2007). “CanSatgeneral introduction and educational advantages”. In *Proceedings of the 18th ESA Symposium on European Rocket and Balloon Programmes and Related Research, Visby*.

[8] Aliyev et al., (2017). "Design of solar powered subscale glider for CanSat competition," 2017 8th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, pp. 453457.

[9] H. Aly et al., (2013). "Project-based space engineering education: Application to autonomous rover-back CanSat," 2013 6th International Conference on Recent Advances in Space Technologies (RAST), Istanbul, pp. 1087-1092.

Estrategias innovadoras en los procesos de enseñanza y de aprendizajes de la programación

Gustavo ASTUDILLO[1], Silvia BAST[1], Pedro WILLGING[1], Darío SEGOVIA[1], Leandro CASTRO[1] , Pamela LUCERO[1] , Martín LOBOS[1] & Juan, DISTEL[1]

[1] Departamento de Matemática/FCEyN/UNLPam
astudillo@exactas.unlpam.edu.ar, silviabast@exactas.unlpam.edu.ar, pedro@exactas.unlpam.edu.ar, dariosegovia2000@gmail.com, leajcastro@gmail.com, plpamelalucero@gmail.com, lobmar146@gmail.com, disteljm@gmail.com

RESUMEN

En la primera etapa del proyecto de investigación "Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática", nos hemos propuesto identificar y categorizar distintas propuestas de enseñanza de la programación, como también, definir una secuencia de aprendizaje reutilizable que incluya actividades, recursos y materiales para el aprendizaje de las nociones básicas de programación.

Para la identificación y categorización de las propuestas, se está llevando adelante la búsqueda de las mismas en distintas bases bibliográficas y repositorios, y definiendo el conjunto de criterios para su categorización.

Se ha definido una secuencia de aprendizaje, que fue implementada en talleres para docentes, lo que permitió la evaluación por pares expertos, y en talleres para alumnas/os de nivel medio y para ingresantes a la Universidad, lo que permitirá evaluar la propuesta en un contexto real y medir el impacto en los aprendizajes.

Palabras clave: gamificación, juegos serios, robótica educativa, aprendizaje basado en problemas, pensamiento computacional, programación

CONTEXTO

El grupo de investigación GrIDIE[1] (Grupo de Investigación y Desarrollo en Innovación Educativa) enfoca, desde 2005, su investigación en tecnologías informáticas aplicadas en educación. Desde 2018, el grupo impulsa dos líneas de investigación: "Aprendizaje de las ciencias con tecnologías educativas" e "Incorporación de estrategias innovadoras en los procesos de enseñanza y de aprendizajes de informática". Las mismas se desarrollan y se financian a través de dos proyectos de investigación con evaluación externa y aprobados por resolución 27/18 CD-FCEyN.

1. INTRODUCCIÓN

En este trabajo presentamos los avances realizados durante el primer año del proyecto de investigación "Incorporación de Estrategias innovadoras en los Procesos de Enseñanza y de Aprendizajes de Informática".

La motivación de esta investigación radica en que en la gran mayoría de los planes de estudio de los niveles primario y medio no se incluye la programación, por lo que cuando un estudiante llega a la universidad, no posee en la mayoría de los casos, ideas previas al respecto y desconoce los conceptos centrales de la misma. Ante este cuadro de situación, los docentes de las asignaturas iniciales de programación se encuentran con muchas incógnitas por resolver: por dónde comenzar, qué paradigma utilizar, qué lenguaje seleccionar, cómo plantear las clases teóricas,

[1] Facultad de Ciencias Exactas y Naturales, UNLPam

cómo organizar la práctica, entre otras cuestiones.

Por otro lado, según el reporte de la Fundación Sadosky (2013), informes de organismos de países desarrollados "coinciden en señalar que el éxito que cada país posea para poder enseñar computación con eficacia estará relacionado directamente con la habilidad de dicho país de poder innovar y competir en los mercados actuales".

A lo anteriormente expuesto debe agregarse también el perfil de los estudiantes actuales, usuarios asiduos de tecnologías. 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.

Teniendo en cuenta los aspectos mencionados, los esfuerzos de esta investigación están dirigidos a definir estrategias que faciliten el proceso de enseñanza-aprendizaje de la programación y que además motiven a los estudiantes.

En primer lugar, nos enfocamos en definir las fuentes de información, realizar búsquedas y categorizar las propuestas de enseñanza de la programación que existen (y están publicadas) en nuestro país. Este proceso enriqueció una secuencia didáctica que el grupo viene desarrollando para el aprendizaje de nociones básicas de programación utilizando robótica educativa.

A continuación, se describen: el proceso de búsqueda bibliográfica, una primera categorización de las propuestas y la secuencia didáctica propuesta por el grupo de investigación.

1.1 Revisión de las propuestas

Para la selección de las propuestas, se decidió la revisión de artículos publicados en revistas y congresos argentinos. Como metodología, se utilizó la definida por Kitchenham (2004), en la misma la autora plantea: (a) definir preguntas de investigación, (b) trazar una estrategia de búsqueda (dónde buscar, con qué palabras claves), (c) establecer criterios de inclusión y exclusión, que serán aplicados, tanto para la selección inicial, como para la selección final.

En un primer análisis de los artículos localizados surge que están enfocados fuertemente en el desarrollo del pensamiento computacional. También, pudo observarse el uso de la gamificación, los juegos serios, el m-learning y la robótica educativa (o combinación de éstos) como estrategias o recursos didácticos en los que se basan las propuestas.

Pensamiento Computacional

"El pensamiento computacional es el proceso de pensamiento involucrado en la formulación de un problema y la expresión de su(s) solución(es) de tal manera que una computadora, humana o mecánica, pueda llevarla(s) a cabo efectivamente" (Wing, 2017, p. 8).

Según Zapata Ros (2015) algunas habilidades propias del pensamiento computacional y que buscamos estimular en los estudiantes son "el análisis ascendente, [...] el pensamiento divergente o lateral, la creatividad, la resolución de problemas, el pensamiento abstracto, la recursividad, la iteración, los métodos por aproximaciones sucesivas, el ensayo-error" (p. 11).

Gamificación

Muchos de nuestros estudiantes, tienen en los videojuegos una fuente informal de formación. Recursos que ya habían sido señalados como de potencial educativo por el Informe Horizonte (Johnson, Levine, & Smith, 2009).

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, Khaled, Nacke, & Dixon, 2011, p. 2). Autores como McGonigal (2011) o Kapp (2012) afirman que las propuestas gamificadas deben establecer: objetivos, reglas, desafíos o conflictos, competencia, colaboración y/o cooperación, retroalimentación y re-jugabilidad (*re-playability*).

Juegos serios

Los juegos serios son juegos digitales diseñados para educar, entrenar o informar (Michael & Chen, 2005) . Existe actualmente, una importante cantidad de este tipo de juegos orientados al aprendizaje de distintos conceptos de informática, particularmente a la programación de computadoras (Astudillo, Bast, & Willging, 2016).

M-learning

Según la UNESCO (West & Vosloo, 2013) el *m-learning* o aprendizaje móvil implica el uso de dispositivos móviles para facilitar el aprendizaje formal e informal en cualquier momento y lugar.

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 misma, "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).

1.2 Nuestra secuencia didáctica

La propuesta diseñada desde el grupo de investigación hace uso de la robótica educativa y se sustenta en los principios del construccionismo**,** el buen aprendizaje y aprendizaje basado en problemas. Para su consecución fue necesaria la definición de kit (sensores y actuadores), la selección de un IDE, un simulador (para tareas extra-clase) y de un conjunto de ejercicios y problemas que hacen uso de los mismos, en el marco de la secuencia de aprendizaje. Recorriendo en ella los conceptos de estructuras de control (secuencia, repetición y selección) y la noción de variable. Asimismo, se desarrollaron un conjunto de *concept cards* que extienden la secuencia en busca de la transferencia de los aprendizajes a partir del planteo de problemas.

El construccionismo, fue propuesto por Seymour Papert (1990), y establece que el conocimiento se construye y lo hace en la cabeza del que aprende. "El construccionismo nos recuerda que la mejor manera de hacerlo es construir algo tangible -algo fuera de tu cabeza- que es también personalmente significativo." (Papert, 1990, p. 14). En este sentido, los estudiantes construyen robots sencillos y se implican con ellos al programarlos.

Se busca con la secuencia lograr "buenos aprendizajes", que a decir de Juan Ignacio Pozo (2008) son aquellos que producen "un cambio duradero y transferible a nuevas situaciones como consecuencia directa de la práctica realizada" (p. 162).

La propuesta está centrada en el trabajo del estudiante y hace uso del problema como disparador para la adquisición de nuevos conocimientos. Ambos aspectos que son propios del aprendizaje basado en problemas (ABP). Barrows (1986) define al ABP como "un método de aprendizaje basado en el principio de usar problemas como punto de partida para la adquisición e integración de los nuevos conocimientos".

En la secuencia se presentan ejercicios, que se resuelven con apoyo del docente (razonamiento experto). La "realización de ejercicios se basa en el uso de destrezas técnicas sobreaprendidas" (Pozo & Pérez Echeverría, 1994, p. 18). Para luego, presentar un problema cuyo objetivo basado en el ejercicio ya realizado. De manera que, para alcanzar ese objetivo, los estudiantes deben partir de los procedimientos o técnicas que conocen (Pozo & Pérez Echeverría, 1994).

2. 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.

3. RESULTADOS OBTENIDOS/ESPERADOS

Entre los avances realizados durante el año 2018 pueden mencionarse:

El relevamiento de más de 65 propuestas que aplican la gamificación, *m-learning*, robótica educativa, juegos serios para el aprendizaje de la programación.

El análisis de las propuestas permitió realizar una primera categorización de las mismas.

Se definieron una secuencia de aprendizaje, un kit de robótica económico y un conjunto de ejercicios y problemas, se identificaron un IDE (*Integrated Development Environment*), un simulador para implementar las soluciones y los materiales y recursos educativos que soportan la secuencia. También, se evaluó su impacto en docentes, estudiantes de nivel medio e ingresantes a la universidad. La retroalimentación de los docentes y estudiantes permitirá su mejora.

En cuanto a los trabajos futuros:

Se espera completar identificación y evaluación de IDEs de programación de robots teniendo en cuenta características de usabilidad de los mismos y los recursos que los mismos incluyen para facilitar el desarrollo de la secuencia de aprendizaje generada dentro del proyecto.

Se recopilará, analizará y evaluará información de herramientas que permitan el desarrollo de conceptos de programación con enfoque lúdico.

Se continuará trabajando en la medición del impacto de la secuencia didáctica propuesta.

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto participan un investigador formado y tres investigadores en formación.

Actualmente, el proyecto cuenta con un estudiante avanzado y dos graduados que se inician en la investigación.

5. BIBLIOGRAFÍA

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.

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011). Gamification: Toward a definition. En *CHI 2011 Gamification Workshop Proceedings*. Vancouver, BC, Canada.

Fundación Sadosky. (2013). CC – 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas. Recuperado de http://www.fundacionsadosky.org.ar/wp-content/uploads/2014/06/cc-2016.pdf

Johnson, L., Adams Becker, S., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *The NMC Horizon Report: Edición Educación Superior 2016*. Austin, Texas: The New Media Consortium.

Johnson, L., Levine, A., & Smith, R. (2009).

Informe Horizonte. Austin, Texas: The New Media Consortium.

Kapp, K. M. (2012). *The gamification of learning and instruction: game-based methods and strategies for training and education*. John Wiley & Sons.

Kitchenham, B. (2004). *Procedures for performing systematic reviews* (Technical Report No. TR/SE-0401) (p. 27). Reino Unido: Keele University. Recuperado de http://tests-zingarelli.googlecode.com/svn-history/r336/trunk/2-Disciplinas/MetodPesquisa/kitchenham_2004.pdf

McGonigal, J. (2011). *Reality is broken: Why games make us better and how they can change the world*. Penguin.

Michael, D. R., & Chen, S. L. (2005). *Serious Games: Games That Educate, Train, and Inform*. Muska & Lipman/Premier-Trade.

Papert, S. (1990). A critique of technocentrism in thinking about the school of the future. Epistemology and Learning Group, MIT Media Laboratory. Recuperado de http://cursa.ihmc.us/rid=1N5PWKQGN-H47Q0R-37ZW/Papert%20critique%20of%20technocentrism.pdf

Pozo, J. I. (2008). Capítulo 4. Los rasgos de un buen aprendizaje. En *Aprendices y maestros: la psicología cognitiva del aprendizaje* (pp. 159-175). Madrid, España: Alianza.

Pozo, J. I., & Pérez Echeverría, M. del P. (1994). Aprender a resolver problemas y resolver problemas para aprender. En *La solución de problemas*. Santillana Madrid.

Prensky, M. (2001). Digital Natives, Digital Immigrants. *On the Horizon*, *9*(5).

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. Recuperado de https://www.unicef.org/argentina/spanish/educacion_Integracion_TIC_sistemas_formacion_docente.pdf

West, M., & Vosloo, S. (2013). *Directrices de la UNESCO para las políticas de aprendizaje móvil*. Francia: UNESCO.

Wing, J. (2017). Computational Thinking's Influence on Research and Education for all. *Italian Journal on Educational Technology*, *25*(2), 7-14.

Zapata-Ros, M. (2015). Pensamiento computacional: Una nueva alfabetización digital. *RED. Revista de educación a distancia*, (46), 1–47.

ESTUDIO DE COMPETENCIAS GENERALES: DE LA UNIVERSIDAD AL ÁMBITO LABORAL

Marcela Dillon [1], Fernando López Gil[2], Russo Claudia[3], Sarobe Mónica[2], David Fernandez[4], Iglesias Pedro[5], Italiano Guido[5]

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, 236-4636945/44

mdillon@advantive.com.ar, {fernando.lopezgil, claudia.russo, monica.sarobe, david.fernandez, pedro.iglesias}@itt.unnoba.edu.ar

RESUMEN

En la actualidad cada vez son más las empresas que evalúan a sus empleados no sólo por el cumplimiento de las tareas, sino por las competencias que estos sean capaces de demostrar en el desempeño de sus funciones.

Sin embargo es importante destacar que las competencias son «variables» nada fáciles de medir, en comparación con los objetivos, a la vez que su desarrollo requiere de una motivación intrínseca más que de una compensación «de fin de año». Por lo tanto, suponen un tratamiento diferente al ser evaluadas, ya que si bien son objetivas, en cuanto a que son observables, también son subjetivas, en cuanto a la percepción que de ellas tiene el examinador.

El objetivo de esta investigación es definir, desarrollar y aplicar métodos de evaluación para las diferentes competencias laborales que se desarrollan desde la universidad a medida que el alumno avanza en el trayecto curricular.

Palabras clave: Competencias básicas, Formación por competencias, Competencias Laborales, Perfiles educativos, Evaluación por competencias.

CONTEXTO

Esta línea de investigación se realiza en el marco del proyecto “Informática y Tecnologías Emergentes", con lugar de trabajo en el

---

1 Docente Investigador externo al ITT

1 Docente Investigador - ITT

1 Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC

1 Becario CIN - ITT

1 Becario de Grado

1 ITT - Centro Asociado CIC

Instituto de Tecnología y Transferencia (ITT) de la Universidad Nacional del Noroeste de la Provincia de Buenos Aires.

El proyecto Perfil.AR comenzó en el año 2013 con el estudio de las competencias básicas en el ámbito de la Escuela de Tecnología como consecuencia de los encuentros Escuela-Empresa que se realizan. En estos encuentros se recogieron inquietudes por parte de las empresas sobre el desarrollo de ciertas competencias por parte de los graduados.

En base a esto se realizaron encuestas para evaluar las necesidades que se planteaban en el mercado laboral y el modo en que desde la universidad se podía ayudar a desarrollarlas. En la actualidad se continúa trabajando en el modo de conciliar las diferentes competencias definidas en el ámbito académico con las que están definidas en el ámbito laboral.

En la realización de este trabajo se integran docentes de grado y posgrado de la UNNOBA, alumnos, becarios e investigadores externos.

1. INTRODUCCIÓN

Los cambios que hoy se producen en el ámbito laboral, caracterizados por la globalización de la economía, y la continua introducción de las nuevas tecnologías en los procesos de producción y administración en las organizaciones, han provocado una constante evolución de los puestos de trabajo, lo cual hace difícil mantener su estabilidad. Esto ha provocado una modificación en el contrato entre las organizaciones y sus miembros, basándose ahora en el desarrollo profesional de los empleados y en la búsqueda, por parte de dichas organizaciones, de nuevas formas para potenciar al máximo las competencias de su personal.

La formación profesional no está solamente en el conocimiento específico de su ámbito de estudio, sino en la formación de conocimientos, habilidades, actitudes y valores que la persona desarrolla y que le permiten desempeñarse de manera apropiada en cualquier entorno productivo, más allá del sector económico, el nivel del cargo, la complejidad de la tarea o el grado de responsabilidad. El concepto de competencia surge de la necesidad de valorar no sólo el conjunto de los conocimientos apropiados (saber) y las habilidades y destrezas (saber hacer) desarrolladas por una persona, sino de apreciar su capacidad de emplearlas para responder a situaciones, resolver problemas y desenvolverse en el mundo. Estamos hablando del desarrollo de capacidades y habilidades que le permiten al profesional, tener una inteligencia práctica y una mentalidad emprendedora. Igualmente, implica una mirada a las condiciones del individuo y disposiciones con las que actúa, es decir, al componente actitudinal y valorativo (saber ser) que incide sobre los resultados de la acción. Se pretende formar sujetos capaces de desempeñarse proactivamente en un mundo globalizado, como miembros útiles de la sociedad. Todo esto es necesario que se desarrolle, siempre considerando el nivel de madurez de la persona, desde el nivel inicial de la escuela y se continúe en la universidad; haciendo natural el desarrollo continuo que deberá seguir en el ámbito profesional y laboral. La universidad en el siglo XXI debe tomar en cuenta estos cambios en las relaciones laborales para formar profesionales que puedan ser capaces de responder a estos nuevos desafíos.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La Agencia Nacional de Evaluación de la Calidad y Acreditación de España (ANECA) ha tomado como base de referencia la propuesta Tuning para la formulación de las competencias transversales de las nuevas titulaciones en las universidades españolas. Diferencian las competencias genéricas y las específicas de la siguiente manera:

- competencias genéricas son aquellas que son compartidas y que pueden generarse en cualquier titulación.
- competencias específicas son aquellas que están asociadas a áreas de conocimiento concretas.

Competentencias instrumentales	*Competencias interpersonales*	*Competencias sistémicas*
1. Capacidad de análisis y síntesis 2. Capacidad de organizar y planificar 3. Conocimientos generales básicos 4. Conocimientos básicos de la profesión 5. Comunicación oral y escrita en la lengua propia 6. Conocimiento de una segunda lengua 7. Habilidades básicas de manejo del ordenador 8. Habilidades de gestión de la información 9. Resolución de problemas 10. Toma de decisiones	11. Capacidad crítica y autocrítica 12. Trabajo en equipo 13. Habilidades interpersonales 14. Capacidad de trabajar en un equipo interdisciplinar 15. Capacidad para comunicarse con expertos de otras áreas 16. Apreciación de la diversidad y de la multiculturalidad 17. Habilidad de trabajar en un contexto internacional 18. Compromiso ético	19. Capacidad de aplicar los conocimientos en la práctica 20. Habilidades de investigación 21. Capacidad de aprender 22. Capacidad de adaptarse a nuevas situaciones 23. Capacidad de generar nuevas ideas (creatividad) 24. Liderazgo 25. Conocimiento de culturas y costumbres de otros países 26. Habilidad para trabajar de forma autónoma 27. Diseño y gestión de proyectos 28. Iniciativa y espíritu emprendedor 29. Preocupación por la calidad 30. Motivación de logro

Gráfico 1 -. Competencias genéricas

En esta misma línea se sitúan numerosos países europeos. Estas competencias básicas genéricas se caracterizan por su potencialidad de transferencia y su multifuncionalidad tanto en las distintas áreas disciplinares como en situaciones de la vida diaria, así como por su durabilidad en el tiempo. Estas competencias genéricas no se aprenden de forma aislada, sino integradas en los distintos escenarios de la vida (educativo, profesional-laboral, comunitario y personal). En el escenario de la enseñanza formal, tienen sentido en la medida que se integran en los contenidos de todas las áreas curriculares. Actualmente las organizaciones tienden a determinar, de acuerdo a las necesidades de su entorno y su estrategia empresarial, cuáles son las competencias que realmente producen un rendimiento superior. Para esto se emplean una serie de técnicas tales como el panel de expertos, las entrevistas, y las entrevistas focalizadas.

Las competencias pueden estudiarse de manera conjunta a partir del análisis de la función directiva. Basándonos en el modelo antropológico de empresa propuesto por Pérez López (1988) La función directiva consiste en diseñar estrategias que produzcan valor económico, desarrollando las capacidades de sus empleados y uniéndolas con la misión de la empresa.

Actualmente las organizaciones tienden a determinar, de acuerdo a las necesidades de su entorno y su estrategia empresarial, cuáles son las competencias que realmente producen un rendimiento superior. Para esto se emplean una serie de técnicas tales como el panel de expertos, las entrevistas, y las entrevistas focalizadas.

En este grupo se acordó aceptar la definición de competencias como …"*procesos complejos de desempeño con idoneidad en determinados contextos, integrando diferentes saberes (saber ser, saber hacer, saber conocer y saber convivir), para realizar actividades y/o resolver problemas con sentido de reto, motivación, flexibilidad, creatividad, comprensión y emprendimiento, dentro de una perspectiva de procesamiento meta-cognitivo, mejoramiento continuo y compromiso ético, con la meta de*

contribuir al desarrollo personal, la construcción y afianzamiento del tejido social, la búsqueda continua del desarrollo económico-empresarial sostenible, y el cuidado y protección del ambiente y de las especies vivas"… para que teniéndolas ya definidas poder avanzar en la evaluación de las mismas.

3. RESULTADOS OBTENIDOS / ESPERADOS

El objetivo general de este proyecto es estudiar la conciliación entre el desarrollo de competencias generales en la educación y el desarrollo de las mismas en el ámbito profesional/laboral de modo de potenciar desde el ámbito académico la inserción laboral de los egresados. Como objetivos específicos podemos encontrar:

- Estudiar la relación entre las competencias generales definidas para la educación y las competencias generales que se requiere en el ámbito laboral:
  - cómo se definen en cada ámbito
  - cómo se traduce cada competencia de un ámbito a otro
  - cómo se desarrollan en cada ámbito
  - cómo se miden en cada ámbito
- Determinar la necesidad de proponer algún cambio que le permita a la persona, un desarrollo armonioso y continuado de las competencias generales.

4. FORMACIÓN DE RECURSOS HUMANOS

Se pretende propiciar un marco formal con el establecimiento de esta línea de investigación, para atender a la formación de recursos humanos, en su rol de investigadores o partícipes activos en el equipo de investigación; fomentando así la culminación de sus estudios superiores, y promoviendo la redacción, exposición y defensa de Trabajos Finales de Grado y Postgrado.

También en el término de tres años, se espera contribuir al inicio y concreción de 1 (una) Tesis de Magíster y la finalización de 1 (una) Tesis Doctoral.

5. BIBLIOGRAFÍA

Inés Parra de Azuero (2010), "Articulación de la Educación con el Mundo Productivo, Competencias Laborales Generales" Colombia.

Claudia Russo, Mónica Sarobe, Marcela Dillon y otros (2015) "El desafío de las nuevas tecnologías de la información y las comunicaciones en los contextos educativos" WICC 2015. Universidad Nacional de Salta.

Tobón, S. (2008). "La formación basada en competencias en la educación superior: el enfoque complejo". México: Universidad Autónoma de Guadalajara.

OCDE (2002): "Définitions et sélection des compétences. Fondements théoriques et conceptuels. Documents de strategie". DEELSA/ED/CERI/CD

Recomendación del Parlamento Europeo y del Consejo, de 18 de diciembre de 2006, sobre las competencias clave para el aprendizaje permanente (2006/962/CE)

TUNING (2003): "Tuning Educational

Structures in Europe. Final Report, Phase One." University of Deusto/University of Groningen. www.relint.deusto.es

Carro, L. (2004): "Los grados de magisterio en los 25 países de la Unión Europea". En: Título de grado de Magisterio. Volumen 2. Agencia Nacional de Evaluación de la Calidad y Acreditación (ANECA)

Garagorri, X. (2007): "Propuestas curriculares basadas en competencias en el ámbito europeo".Aula de Innovación Educativa, nº. 161

Rey, B. (1996). "Les compétences transversales en question". Paris: ESF

Pablo Beneitone y otros (2007) "Reflexiones y perspectivas en Educación Superior en América Latina. Informe Final – Proyecto Tuning LA". Universidad de Deusto – Universidad de Groningen. España.

José Lino Contreras Véliz y otros (2013) "Educación Superior en América Latina: reflexiones y perspectivas en Informática." Universidad de Deusto. España.

Claudia Russo, Mónica Sarobe, Marcela Dillon y otros (2015) "Desarrollo de perfiles profesionales en base a competencias Perfil.AR" TE&ET 2015 Universidad Nacional del Nordeste.

Formación docente y estrategias de enseñanzas innovadoras para fomentar el pensamiento computacional

Gladys N.Dapozo. Cristina L. Greiner, Raquel H.Petris, María C. Espíndola, Ana M. Company, Yanina Medina
Departamento de Informática. Facultad de Ciencias Exactas y Naturales y Agrimensura
Universidad Nacional del Nordeste. 9 de Julio Nº 1449. Corrientes. Argentina.
{gndapozo, cgreiner, rpetris}@exa.unne.edu.ar, mcespindola@yahoo.com, anamacom@hotmail.com, yaninamedinaa@gmail.com

RESUMEN

Este proyecto de investigación se orienta al estudio y aplicación experimental de estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación, y a las didácticas específicas para la formación de los profesores de los niveles educativos preuniversitarios, con el objetivo de contribuir con las políticas públicas orientadas a incorporar las Ciencias de la Computación en las escuelas. El objetivo es contribuir con la promoción del pensamiento computacional e impactar en un mayor desarrollo de carreras STEM (Science, Technology, Engineering and Mathematics).

Palabras clave: Enseñanza de la programación. Didácticas específicas. Pensamiento computacional.

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

Diversos países se ven afectados por una marcada disminución de estudiantes en carreras TIC. En la Argentina, la Fundación Sadosky [1] tiene entre sus objetivos la incorporación del estudio de programación en las escuelas y fomentar el incremento de la matrícula en carreras relacionadas con las TIC. Específicamente, la iniciativa Program.AR planea llevar la enseñanza y el aprendizaje de las Ciencias de la Computación a la escuela, comenzando con la difusión y popularización de la disciplina, la generación de contenidos escolares y la formación docente, entre otros.

En la última década, la web 2.0, los dispositivos móviles, la poderosa industria del videojuego y el resurgir del movimiento DIY (*Do It Yourself*) han puesto en un primer plano las necesidades y carencias en la formación sobre pensamiento computacional y programación. Como respuesta a esta demanda, gobiernos, empresas e instituciones educativas, han desarrollado proyectos y adoptado decisiones relacionadas con el fomento de la programación en la educación [2].

El concepto de pensamiento computacional es una competencia compleja de alto nivel relacionada con un modelo de conceptualización específica de los seres humanos que desarrollan ideas y está vinculada con el pensamiento abstracto-matemático y con el pragmático-ingenieril que se aplica en múltiples aspectos de la vida diaria. Es una competencia básica que todo ciudadano debería poseer para desenvolverse en la sociedad digital. Es una forma de resolver problemas de manera inteligente e imaginativa. Además, posee las características de combinar abstracción y pragmatismo, puesto que se fundamenta en las

matemáticas, un mundo de ideas, y se desarrolla a partir de proyectos de ingeniería que interactúan con el mundo real [2]. El pensamiento computacional refuerza los estándares educativos en todas las asignaturas para acrecentar la habilidad del estudiante de solucionar problemas y así desarrollar pensamiento de orden superior [3]. Este pensamiento está relacionado con otros pensamientos tales como: pensamiento lógico, pensamiento analítico, pensamiento algorítmico, pensamiento abstracto, pensamiento divergente y pensamiento crítico [4].

La sociedad actual y los sistemas de producción, de servicios y de consumo demandan profesionales cualificados en las industrias de la información. Particularmente en el mundo desarrollado se da la paradoja de escasez de profesionales formados en carreras vinculadas con la Informática. Ante esta situación los sistemas educativos de los países más sensibles han abordado el problema desde la perspectiva de una reorganización del curriculum en la mayor parte de los casos donde se ha producido esa reacción. Sin embargo la cuestión de fondo supone la aparición de unas nuevas destrezas básicas. Se trata de una nueva alfabetización, la alfabetización digital, y que como tal hay que comenzar desde las primeras etapas del desarrollo individual, al igual como sucede con otras habilidades claves: en la lectura, la escritura y las matemáticas. El planteamiento consiste en favorecer el aprendizaje de la programación en los niños, de forma progresiva, con tareas desde las más sencillas y más lúdicas a las más complejas [5].

En la UNNE, se decidió acompañar la iniciativa de la Fundación Sadosky de promover el estudio de carreras vinculadas con la Informática, y en este marco, indagar sobre el perfil de los alumnos del nivel medio en relación a su formación y habilidades tecnológicas, su interés por la elección de carreras vinculadas con la Informática y los factores que influyen en esta elección, a fin de aportar información que contribuya al objetivo de acercar a los jóvenes a la Informática, como actividad profesional. Así también, ha realizado capacitación a los docentes de distintos niveles educativos. En estas actividades se han recabado datos con el propósito de aportar información a la problemática de las vocaciones TIC y la promoción e incorporación del pensamiento computacional en los niveles educativos no universitarios.

La Iniciativa Program.AR sostiene que el mundo moderno no puede comprenderse cabalmente sin contar con un conjunto de herramientas que permitan decodificar la lógica de la tecnología que media en buena parte los vínculos que establecemos con el mundo en el que estamos insertos. Los diseñadores de esta propuesta sostienen que la lógica de funcionamiento de la tecnología no se revela con su mero uso [6]. Con esto quieren señalar que hay una gran diferencia entre una persona en condiciones de utilizar de manera hábil una determinada tecnología y una persona que comprende el funcionamiento de la tecnología que utilice. El objetivo es explorar distintos aspectos de las Ciencias de la Computación. La programación es una de las áreas más importantes de las Ciencias de la Computación. Como disciplina, la programación está orientada al desarrollo de una serie de habilidades de abstracción y operacionalidad. El primer tipo de habilidades incluye técnicas como la simplificación de problemas, la definición de soluciones generales aplicables a problemas similares y la asignación de nombres significativos a las distintas partes de una solución. El segundo tipo de habilidades, las operacionales, supone la definición de soluciones en términos de un conjunto de pasos que deben ejecutarse en un

orden determinado para alcanzar un objetivo.

En esta línea de acción, el Ministerio de Educación y Deportes de la Nación diseñó el Plan Nacional Integral de Educación Digital (PLANIED) [7], con el propósito de integrar la comunidad educativa en la cultura digital, favoreciendo la innovación pedagógica, la calidad educativa y la inclusión socioeducativa. Considera competencias fundamentales para facilitar la inclusión de los alumnos en la cultura digital las siguientes: Creatividad e Innovación, Comunicación y Colaboración, Pensamiento Crítico y Uso Autónomo de las TIC.

Scratch

Una herramienta muy utilizada para desarrollar el pensamiento computacional es Scratch. Es un lenguaje de programación basado en bloques, donde la gramática visual de los mismos y sus reglas de combinación tienen el mismo rol que la sintaxis en los lenguajes basados en texto como C, Java o Python. Su diseño original surgió en torno a las necesidades e intereses de un grupo de jóvenes que participaban en clubes extra-escolares, y ofrece la posibilidad de programar a través de la exploración y la colaboración entre pares, lo cual fue un aspecto innovador destacable y potencialmente favorable para el desarrollo cognitivo en un marco de aprendizaje activo. Algunos estudios señalan el posible impacto de Scratch en el desarrollo cognitivo, en aspectos como la resolución de problemas, el pensamiento lógico, la motivación, etc. [8]. En este marco, saber programar implica estar alfabetizado en el "código del futuro", entendiendo a la tecnología como un elemento que habilita nuevas formas de pensamiento y acción.

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 primaras 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".

3. RESULTADOS OBTENIDOS/ESPERADOS

Las experiencias surgidas del contacto con docentes que han incorporado la programación en sus escuelas, generadas en las actividades de capacitación fueron expuestas en [9], En este evento fue seleccionado como uno de los dos Mejores Trabajos presentados en relación al interés de las jornadas.

Por otro lado, considerando la motivación que se detectó en los docentes que se han capacitado durante los tres últimos años, la necesidad de avanzar en contenidos y generar proyectos educativos institucionales que permitan abordar de manera sistematizada el Pensamiento Computacional en los niveles no universitarios de la ciudad de Corrientes, se ha creado en al ámbito de la Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste (FaCENA – UNNE) la "Diplomatura en Programación y

Robótica Educativa", por Res Nº 1060/18 CD de 200 horas de duración comprendidas en cinco módulos, cuyo dictado se llevará a cabo en el período lectivo 2019.

El objetivo principal es acompañar las políticas públicas orientadas a incorporar en la educación básica conceptos y herramientas para lograr ciudadanos activos y críticos de este mundo tecnológicamente intensivo, mediante la formación de los docentes del nivel primario, secundario y terciario en el manejo de tecnologías de programación y robótica, adecuadas a los niveles preuniversitarios.

La diplomatura está constituida por 5 módulos de 40 hs. cada uno (que se muestran en la tabla 1), dictados en modalidad presencial con actividades de apoyo a la presencialidad a través de la plataforma de UNNE Virtual.

Tabla 1: Módulo de la Diplomatura en Programación y Robótica Educativa.

Programación visual basada en bloques	40 hs.
Programación de placas Arduino	40 hs.
Fundamentos de Programación	40 hs.
Programación de microcontroladores	40 hs.
Enfoques didácticos para la programación	40 hs.

Como resultado de esta propuesta formativa se espera que los participantes (en este caso, el docente de las escuelas), adquiera las siguientes capacidades y competencias:

- Comprender la importancia de la incorporación en las escuelas de los temas propios de las Ciencias de la Computación.
- Entender y aplicar las técnicas básicas de programación y de robótica para desenvolverse adecuadamente en los nuevos contextos educativos.
- Conocer y aplicar un método de resolución de problemas que facilitará la enseñanza de la programación, y que puede ser extendido a otros ámbitos.
- Aplicar e integrar nuevos enfoques didácticos que promueven el aprendizaje activo y colaborativo.
- Diseñar actividades de enseñanza utilizando dispositivos innovadores que ofrecen retos y desafíos motivadores para los estudiantes.

Se destaca que el curso enfatiza habilidades digitales a través de intensa práctica de programación y robótica, pero en un contexto en el que se tiene en cuenta permanentemente el enfoque didáctico, buscando no solo que el docente aprenda los temas disciplinares sino también, y muy importante, que los sepa transmitir adecuadamente.

Por ello, las actividades de evaluación, además de las cuestiones específicas de cada módulo, evaluaran la adquisición de las competencias digitales propuestas por PLANIED: Creatividad e innovación, Comunicación y colaboración, Pensamiento crítico y Uso autónomo de las TIC.

En cuanto a Resultados Esperados, el proyecto continuará con las siguientes actividades:

a) 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 las actividades de capacitación que el equipo docente realiza mediante convenio con la Fundación Sadosky, y de las actividades de la Diplomatura en Programación y Robótica Educativa, mencionada anteriormente, se evaluarán intereses, expectativas y posibilidades reales de incorporación de contenidos de Ciencias de la Computación en las escuelas. 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.

b) 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.

En la línea vinculada con estrategias educativas que incorporen métodos y herramientas innovadoras para la enseñanza de la programación en las carreras de Informática, se proponen las siguientes actividades:

- Relevar y evaluar herramientas software 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 de las mismas.
- Diseñar y validar un marco de referencia que oficie de guía en la selección de la herramienta más apropiada para el logro de los objetivos de aprendizaje

4. FORMACIÓN DE RECURSOS HUMANOS

En este proyecto están involucrados seis docentes investigadores y dos tesistas de posgrado desarrollan su trabajo final de la Maestría en Tecnologías de la Información de la UNNE.

5. REFERENCIAS

[1] Fundación Sadosky. Inestigación y desarrollo en TIC. http://www.fundacionsadosky.org.ar/

[2] Jesús Valverde Berrocoso, María Rosa Fernández Sánchez, María del Carmen Garrido Arroyo. El pensamiento computacional y las nuevas ecologías del aprendizaje. RED-Revista de Educación a Distancia, 46(3). 15-Sept.-2015 DOI: 10.6018/red/46/3. Disponible en: http://www.um.es/ead/red/46/valverde_et_al.pdf

[3] Computer Science Teachers Associations (CSTA) and International Society for Technology in Educaction (ISTE) (2011). Computational thinking teacher resources (2nd ed.). Recuperado el 14 de Marzo de 2019, de https://id.iste.org/docs/ct-documents/ct-teacher-resources_2ed-pdf.pdf?sfvrsn=2

[4] IKER CORTÉS / MADRID Recuperado 14 de Marzo de 2019, de https://www.abc.es/20121016/tecnologia/rc-poder-educativo-videojuego-201210161402.html

[5] Miguel Zapata-Ros. Pensamiento computacional: Una nueva alfabetización digital. RED-Revista de Educación a Distancia, 46(4). 15-Sept.-2015 DOI: 10.6018/red/46/4

[6] Pablo Factorovich Federico Sawady O'Connor. Actividades para aprender a Program.AR. Volumen 1. Edición: Ignacio Miller. E-Book. - ISBN 978-987-27416-1-7.

[7] Ministerio de Educación y Deportes de la Nación diseñó el Plan Nacional Integral de Educación Digital (PLANIED) http://planied.educ.ar/

[8] Natalia Monjelat, Patricia Silvana San Martín. Programar con Scratch en contextos educativos: ¿asimilar directrices o co-construir tecnologías para la inclusión social? Praxis Educativa. VOL. 20, NÚM. 1 (2016) .ISSN: 0328-9702.

[9] Gladys Dapozo, Raquel H. Petris, Cristina L. Greiner, Ana M. Company, María C. Espíndola "Formación docente para incorporar la programación en las escuelas" Vol 3 (2018) Actas de las XXIV Jornadas sobre Enseñanza Universitaria de la Informática – JENUI - Barcelona España, pp. 23-30. Disponible en http://actasjenui.aenui.net/

La tecnología como vehículo de articulación Nivel Medio / Universidad

Fernanda Beatriz Carmona, Alberto Eduardo Riba, Fernando Emmanuel Fratti, Claudio Barrionuevo, Matías Pérez, Emmanuel Alejandro Portugal, José Nicolás Frati, Iván Flores, Pablo Oporto Poblete, Andrés Nicolás Ortiz

Departamento de Básicas y Aplicadas, Universidad Nacional de Chilecito
9 de Julio 22, Chilecito, La Rioja, Argentina
{fbcarmona, ariba, fefrati, cbarrionuevo}@undec.edu.ar, {mataguper20, emmanuel.portugal.91, nicofrati, ivanflores844, pablogabrieloporto, andresnicolasortiz}@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 Nº EXP-S01:0001042/2016, período 2016-2017 "Universidad, Cultura y Sociedad", SPU y su continuidad 2017-2018 aprobada por RESOL-2017-5135-APN-SECPU, 2018-2019 EU43-UNDEC13713 convocatoria 2018, en proceso de evaluación.

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 Nº 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 impreso 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. En el 2018 se actualizo la plataforma para facilitar su uso como instrumento de soporte a las clases prácticas de las asignaturas involucradas, permitiéndose la creación de colecciones de problemas seleccionados por los docentes y adecuar la corrección de problemas para permitir inspecciones de código antes de dar por correcto un problema corregido por el ayudante virtual. Se desarrollaron talleres semanales orientado a jóvenes de nivel secundario, en tres instituciones de nivel medio, que incluyeron nociones básicas sobre programación y uso de la plataforma de aprendizaje colaborativo, ajustadas a los conocimientos previos de los asistentes. Estos talleres fueron coordinados por los docentes, pero desarrolladas por los estudiantes universitarios. 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 /2017/2018 con la participación de alumnos de grado.

Formación de Recursos Humanos

El equipo de investigación de esta línea de trabajo está compuesto por cuatro docentes, dos egresados y cuatro alumnos de grado de la Universidad Nacional de Chilecito. 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, dos auxiliares docentes y dos egresados de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC. También participan 5 alumnos avanzados de grado.
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. ttp://www.unidiversidad.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] Josfie Joaqufifin Brunner and Rocío Ferrada Hurtado, “Educación superior en Iberoamerica”, Eds. Santiago, Chile: RIL R editores, Oct. 2011.

[7] SPU, “Articulación escuela secundaria educación superior," Blog: Portal de Educación, 2007.

[8] Ana Lourdes Acuña, Marfiá 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 Inov. 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.

Las Ciencias de la Computación en el Currículum Escolar

Gisela Branchini[2] Marcos Manuel Cortez[2]
María de los Angeles Pedemonte[2] Jorge Rodríguez[1]

gbranchini@gmail.com, mmcortez@neuquen.gov.ar,totipedemonte@yahoo.com, j.rodrig@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] *Dirección Provincial Educación Secundaria*
Consejo Provincial de Educación
Ministerio de Educación de la Provincia de Neuquén

Resumen

En la última década la enseñanza de los conceptos sobre Ciencias de la Computación irrumpe con fuerza en las instituciones escolares, contando con gran consenso y siendo reconocida como disciplina prioritaria, tanto en países desarrollados como en desarrollo.

Argentina se encuentra articulando diversas acciones que proponen la implementación de estrategias para aproximar las Ciencias de la Computación a la Escuela Secundaria.

En este contexto en 2018 la provincia de Neuquén incorpora la asignatura Informática al Ciclo Básico Común para la educación secundaria con un fuerte anclaje a las Ciencias de la Computación.

Sin embargo, en el mapa curricular nacional las Ciencias de la Computación tienen una distribución desigual subsumida, generalmente, en el área de Tecnología.

Se propone analizar cuáles son los Paradigmas y Enfoques con que la computación se representa en los documentos curriculares.

La línea de investigación busca identificar paradigmas y enfoques que logren constituirse en categorías de análisis satisfactorias para explicar la situación de la computación en los diseños curriculares.

La relevancia de este estudio radica en la necesidad de disponer de herramientas que permitan mejorar la comprensión del sentido pedagógico con el que la computación se expresa en las propuestas de enseñanza oficiales para el nivel secundario.

Palabras Clave: Ciencias de la Computación, Currículum, Escuela Secundaria, Enseñanza de la Computación.

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.

El convenio tiene como principal objetivo contribuir recíprocamente al desarrollo de actividades de investigación, formación de recursos humanos y promoción de la enseñanza de las Ciencias de la Computación durante la escolaridad obligatoria.

Este trabajo se desarrolla en el ámbito del proyecto de investigación *Agentes Inteligen-*

tes. 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 comprensión de la naturaleza de los cambios tecnológicos ha conducido a que, actualmente, las Ciencias de la Computación ocupen un lugar relevante tanto por sus aplicaciones prácticas como por introducir una serie de conceptos que operan de un modo transversal en distintas áreas del conocimiento y la cultura.

Esto ha motivado que en la última década la enseñanza de los conceptos de las Ciencias de la Computación irrumpa con fuerza en las instituciones escolares, contando con gran consenso y siendo reconocida como disciplina prioritaria, tanto en países desarrollados como en desarrollo; involucrando a gobiernos, organizaciones civiles, grupos de investigación y docentes de todos los niveles de los sistemas de educación [10, 12].

En el plano nacional, Argentina se encuentra articulando diversas acciones que proponen la implementación de estrategias orientadas a aproximar las Ciencias de la Computación a la Escuela Secundaria vinculando a las Universidades Nacionales, programas nacionales como Program.ar y Aprender Conectados y la Fundación Sadosky.

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 [2].

Es así que en 2016 la Nueva Escuela Secundaria de la Ciudad Autónoma de Buenos Aires modifica los contenidos de la asignatura Tecnologías de la Información, perteneciente al corpus de la Formación General del Ciclo Orientado de todos los bachilleratos incorporando contenidos específicos de las Ciencias de la Computación como son la programación y el pensamiento computacional.

En 2018 el Consejo Federal de Educación aprueba los Núcleos de Aprendizajes Prioritarios para Educación Digital, Programación y Robótica a implementarse en los niveles Inicial, Primario y Secundario del sistema educativo con el objetivo que los estudiantes puedan integrarse plenamente en la cultura digital [3].

En este sentido, en octubre de 2018 la provincia de Neuquén incorpora la asignatura Informática al Ciclo Básico Común en todas las modalidades de la enseñanza secundaria con un fuerte anclaje a las Ciencias de la Computación en el documento curricular del nivel [4].

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 [14, 13].

No obstante, la publicación de informes que dan cuenta de la concreción de distintas y diversas iniciativas no reflejan fehacientemente el fenómeno en que se encuentran incorporadas las Ciencias de la Computación en la educación secundaria [8, 7, 15].

Se entiende al currículum escolar como un proyecto cultural, educativo y una herramienta teórico metodológica del trabajo docente que supone una respuesta a las preguntas del qué, cómo, por qué y para qué enseñar. 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].

Así mismo, el currículum es un campo político de debates y conflictos. Es un terreno de tensiones sobre los intereses de distintos grupos en proponer sus ideas oficialmente válidas. En él encontramos elementos y categorías de análisis para comprender el

complejo entramado histórico y las relaciones de poder que subyacen [6].

Por lo tanto se aprecia que en el mapa curricular nacional las Ciencias de la Computación tienen una distribución desigual subsumida, generalmente, en el espacio correspondiente al área de Tecnología. 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.

Se propone analizar cuáles son los Paradigmas y Enfoques con que la computación se representa en los documentos curriculares.

En este marco se plantea el estudio de los objetivos y propósitos de las Ciencias de la Computación en el currículum escolar con la intención de aportar elementos que colaboren en los procesos de discusión.

2. Línea de Investigación

El estudio a desarrollar en esta Línea de Investigación consiste en identificar paradigmas y enfoques curriculares que permitan describir con rigurosidad la situación de la computación en las propuestas de enseñanza destinadas a la escuela secundaria.

Para ello se propone desarrollar la siguiente línea de trabajo:

- Paradigmas y Enfoques de las Ciencias de la Computación en el currículum escolar

  Un paradigma establece aquello que debe ser observado; la clase de interrogantes que deben desarrollarse para obtener respuestas en torno al propósito que se persigue; qué estructura deben poseer dichos interrogantes y marca pautas que indican el camino de interpretación para los resultados obtenidos de una investigación de carácter científico.

  Desde el punto de vista epistemológico se establece entonces un concepto de realidad basado en el método experimental que fragmenta los conceptos y los problemas de que se ocupa para reorganizarlos después de acuerdo a una cierta lógica.

  En las ciencias sociales, el paradigma está relacionado con el concepto de cosmovisión. Las cosmovisiones son el conjunto de experiencias, creencias y valores que conforman la imagen o figura general del mundo que tiene una persona, a partir de las cuales interpreta su propia naturaleza y la de todo lo existente en el mundo. Una cosmovisión define nociones comunes que se aplican a todos los campos de la vida, desde la política, la economía o la educación como en este caso.

  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*[5]. Esta propuesta constituye la base para las definiciones adoptadas en este trabajo para el cual se identifican los siguientes paradigmas:

  Alfabetización Digital, plantea el desarrollo de competencias digitales básicas,es decir, el conjunto de habilidades para usar satisfactoriamente las TIC [5, 1].

  Competencias TIC, considera la construcción y empleo de estrategias utilizando sistemas informáticos preexistentes para satisfacer necesidades relacionadas a campos específicos como la industria, el comercio o el arte [5, 9].

  Mejoramiento de la Calidad de los Aprendizajes, contempla la introducción transversal de la computación como un

medio para favorecer la construcción de conocimientos en otras disciplinas [1, 9].

Ciencias de la Computación, propone considerar pertinente la inclusión de principios fundamentales de la disciplina científica en los documentos curriculares para el nivel secundario [5, 10].

Cada uno de estos paradigmas presentan diferentes posturas ya que tratan de explicar aspectos como la concepción de la educación, los objetivos de la enseñanza y la organización curricular, entre otros, estableciendo sus parámetros de legitimación del conocimiento.

Es así como en el marco del paradigma de *Ciencias de la Computación* en esta Línea de Investigación se identifican los siguientes enfoques:

Económico Productivo, contempla el prestar atención a las demandas formativas expresadas por la industria y a los productos utilizados en este ámbito.

Del Pensamiento Computacional, centrado en las estrategias de abordaje de problemas transferibles a otros campos problemáticos.

Conocer sobre Ciencias de la Computación para Comprender el Mundo, enfocado en identificar los conocimientos computacionales que favorecen la compresión del mundo.

Relacionado al Acceso al Conocimiento, contempla el prestar atención a los conceptos fundamentales de la disciplina como dimensión significativa en el proceso de organización y consolidación de una ciudadanía plena en derechos. Se hace énfasis en la rigurosidad disciplinar y se propone un recorrido amplio por las áreas de conocimiento de la misma.

La línea de investigación busca establecer si los paradigmas y enfoques enunciados en esta propuesta constituyen categorías de análisis satisfactorias para explicar la situación de la computación en los diseños curriculares.

La relevancia de este estudio radica en la necesidad de disponer de herramientas que permitan mejorar la comprensión del sentido pedagógico con el que la computación se expresa en las propuestas de enseñanza oficiales para el nivel secundario.

3. Resultados

Inicialmente, se realizó un estudio sobre documentos que analizan la situación de las Ciencias de la Computación en propuestas curriculares en Estados Unidos y Europa.

Este trabajo buscó construir definiciones preliminares acerca de los paradigmas y enfoques, articulando aproximaciones documentadas en fuentes diversas. Estas construcciones son presentadas en la sección Línea de Investigación de este artículo.

Es importante mencionar que los paradigmas y enfoques aparecen en forma dispersa y con bajo grado de articulación en la bibliografía consultada. Esta situación dificulta la comprensión de la situación de la computación en el currículum escolar.

En el contexto de esta Línea de Investigación se espera obtener los siguientes resultados:

- Avanzar en la definición de paradigmas como categorías de análisis
- Avanzar en la definición de enfoques como categorías de análisis
- Aportar elementos que colaboren con el análisis de la situación de la computación en las propuestas de enseñanza.

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 dos de los autores de este artículo cursan maestrías orientadas a conocer, comprender y analizar procesos relacionados con las tecnologías en la educación.

Referencias

[1] CEPAL, OEI, and S. G. Iberoamericana. *Metas educativas 2021*. Organización de Estados Iberoamericanospara la Educación, la Ciencia y la Cultura (OEI), 2010.

[2] Consejo Federal de Educación. Res 263/15. *Resoluciones CFE*, 2015.

[3] Consejo Federal de Educación. Res 343/18. *Resoluciones CFE*, 2018.

[4] Consejo Provincial de Educación. Res 1463/18. *Resoluciones CPE*, 2018.

[5] S. Furber. *Shut down or restart? The way forward for computing in UK schools* . The Royal Society Education Section, 2012.

[6] H. A. Giroux. Teachers, public life, and curriculum reform. *Peabody Journal of Education*, 69(3):35–47, 1994.

[7] Google and Gallup. *Searching for computer science: Access and barriers in U.S. K–12 education*. 2015.

[8] Google and Gallup. Trends in the state of computer science in u.s. k-12 schools. 2016.

[9] ISTE. ISTE standards for students". *ISTE Standards*, 2016.

[10] K-12 Computer Science Framework Steering Committee. *The K–12 Computer Science Framework*. ACM, 2016.

[11] Ley Nacional. 26.206 Ley de Educación Nacional. *Disposiciones Generales*, 2006.

[12] F. Sadosky. *CC – 2016 Una propuesta para refundar la enseñanza de la computación en las escuelas Argentinas* . Fundación Sadosky, Argentina, 2013.

[13] 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.

[14] S. Sommer, J. E. Sznek, and J. Rodríguez. Divulgando temáticas computacionales-internet segura. In *X Congreso sobre Tecnología en Educación & Educación en Tecnología*, 2015.

[15] M. Webb, N. Davis, T. Bell, Y. J. Katz, N. Reynolds, D. P. Chambers, and M. M. Sysło. Computer science in k-12 school curricula of the 2lst century: Why, what and when? *Education and Information Technologies*, 22(2):445–468, 2017.

Los Entornos Virtuales Flexibles para el Desarrollo de Competencias en el Área de Algoritmos y Lenguajes de Programación

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 0264-4234980Dirección, Teléfonos
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 de la enseñanza - 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 integren 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.

Palabras clave:
Aprendizaje de Programación. Educación Formal-Informal. Entorno Virtual Flexible. Desarrollo de competencias informáticas.

CONTEXTO

Proyecto de Investigación "*Los entornos virtuales flexibles para el desarrollo de competencias en el área de algoritmos y lenguajes de programación*" Código 21/E1113, Vigentes en el periodo 2018-2019, de la Universidad Nacional de San Juan. Argentina CICITCA. Resolución N°21/18-CS. Este proyecto forma parte del trabajo conjunto del Programa permanente de Investigación del Gabinete de Tecnología e Innovación Educativa (GATIE) y el Departamento del Informática.

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 (Garcia-Holgado & Garcia-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 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.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes a abordar en esta investigación son, el conocimiento tecnológico, pedagógico, disciplinar y de teorías psico-sociales, 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.

las competencias específicas para el aprendizaje de la Programación

Se analizó información del Proyecto Alfin-EEES[1] y Proyecto Alfa Tuning[2] y competencias específicas de Informática. A partir de esto se realizaron talleres para definir con cuales se trabajaría, para lo cual nos enfocamos en competencias genéricas relacionadas con la gestión de la información en categorías relacionadas con:

- Aprender a aprender
- Aprender a buscar y evaluar la información
- Aprender a analizar y sintetizar
- Aprender a generar conocimiento
- Aprender a trabajar juntos
- Usar tecnología para aprender

A partir de la información recopilada, se propusieron estrategias de enseñanza que se aplicaron con alumnos de primer año de la cátedra Procedural de las carreras Licenciatura en Sistemas de Información y Licenciatura en Ciencias de la Computación, del Departamento de Informática FCEFyN - UNSJ. Con ellos se aplicaron las herramientas Plickers, Socrative y

[1] http://www.mariapinto.es/alfineees/contenido.htm
[2] http://www.tuningal.org/es/competencias/especificas/informatica

CodeBoard, registrándose la experiencia para aportar a la evaluación de estrategias.

También se dictó un curso de Lenguaje Phyton destinado a estudiantes de las carreras del departamento de Astronomía y Geofísica. En el mismo se aplicaron herramientas digitales como Jupyter notebook, Padlet, Vokoscree, Audacity, DU-Recorder.

En ambas situaciones las experiencias permitieron un visible acercamiento al logro de los objetivos propuestos.

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 Nº 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? eLearningPapers*, volumen (2(1)), pp(1-7). http://www.elearningeuropa.info/files/media/media11561.pdf

BlasAritio, 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/las-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).

ReigDolors (2017*) Blog El Caparazón referido al ámbito de la innovación, la educación y la tecnología* http://www.dreig.eu/caparazon/

Tejada DìazRafael (2017) *Paradigmas Prospectivos de los contextos Univeristarios*.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.

Programación por Bloques en cátedras de Ingeniería en Informática

Hernán C. Ahumada[1], Daniel A. Rivas[1], Oscar E. Quinteros[1], Cecilia E. Gallardo[1], Nelson A. Contreras[1], Marta del V. Miranda[1], Javier Favore[1], María V. Póliche[1]

[1]Fac. de Tecnología y Ciencias Aplicadas - Universidad Nacional de Catamarca

Maximio Victoria 55, 4700 Catamarca, Argentina

{hcahumada, darivas, oequinteros, ceciliagallardo, nacontreras, mvmiranda, jfavore, vpoliche}@tecno.unca.edu.ar

Resumen

Los datos de rendimiento académico de la Facultad de Tecnología y Ciencias Aplicadas indican que la materia Programación I es la que mayor cantidad de cursadas requiere para alcanzar la condición de alumno regular en la carrera Ingeniería en Informática. Tal situación da cuenta de la dificultad de los alumnos para desarrollar las capacidades básicas del Pensamiento Computacional y de los fundamentos de la Programación. Para revertir, en el corto y mediano plazo, la problemática expuesta se implementaron prácticas de enseñanza alternativas incorporando como recurso didáctico algunas herramientas de programación por bloques.

Palabras clave: Pensamiento Computacional, Programación por bloques.

Contexto

Esta línea de trabajo se viene llevando a cabo en el marco de un proyecto de investigación acreditado por SECyT - UNCA dentro del Programa de Incentivos mediante Resolución rectoral 375/2018, denominado *"Estudio del impacto de enseñar mediante técnicas de programación por bloques las nociones básicas de programación a alumnos de las cátedras Fundamentos de Informática (1º año) y Programación I (2º año) de la carrera Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas (FTyCa) de la Universidad Nacional de Catamarca (UNCa"),* que se desarrolla en el Departamento de Informática de la Facultad de Tecnología y Ciencias Aplicadas (Universidad Nacional de Catamarca).

Introducción

Wing (2006) expresa que "el Pensamiento Computacional es el proceso de pensamiento involucrado en la formulación de problemas y expresión de las soluciones en una forma que pueda ser efectivamente llevada a cabo por un agente de procesamiento de información".

Para Olabe (2015), "el Pensamiento Computacional es una metodología basada en la implementación de los conceptos básicos de las ciencias de la computación para resolver problemas cotidianos, diseñar sistemas domésticos y realizar tareas rutinarias".

La Sociedad Internacional para la Tecnología en Educación (ISTE) y la

Asociación de Docentes en Ciencias de la Computación (CSTA) (ISTE and CSTA, 2011), establecen una definición operativa del Pensamiento Computacional (PC) como un proceso de solución de problemas que incluye, entre otras, las siguientes características:

- Formular problemas de manera que permitan usar computadoras y otras herramientas para solucionarlos.
- Organizar datos de manera lógica y analizarlos.
- Representar datos mediante abstracciones, como modelos y simulaciones.
- Automatizar soluciones mediante pensamiento algorítmico (una serie de pasos ordenados.
- Identificar, analizar e implementar posibles soluciones con el objeto de encontrar la combinación de pasos y recursos más eficiente y efectiva.
- Generalizar y transferir ese proceso de solución de problemas a una gran diversidad de éstos.

Según Brennan y Resnick (2012) el Pensamiento Computacional involucra tres dimensiones en las que se puede evaluar si el estudiante adquiere la capacidad de aplicarlas en la resolución de problemas.

Dimensión 1: Conceptos Computacionales. Son aquellos conceptos que permiten implementar la estrategia de solución a problemas. Tales conceptos son: comandos, secuencias de comandos, procedimientos, repetición simple, alternativa condicional, repetición condicional, sensores, variables, parametrización e interactividad.

Dimensión 2: Prácticas Computacionales. Se refiere a la habilidad de aplicar los conceptos computacionales siguiendo criterios de buenas prácticas en Programación. Entre esas prácticas computacionales están: la descomposición en subproblemas, utilización de procedimientos, legibilidad del programa, reutilización.

Dimensión 3: Perspectivas Computacionales. Perspectiva de expresar la estrategia de solución, perspectiva de trabajar colaborativamente para encontrar la solución, perspectiva de preguntar si es posible dar solución a otros problemas desde un punto de vista computacional.

En años recientes se han presentado varias herramientas para motivar y facilitar el aprendizaje inicial en programación a niños y jóvenes. La gran mayoría de ellas adoptan el enfoque de programación por bloques. Entre las más difundidas se encuentran Scratch (Resnick et. al, 2009) y Alice (Dann et. al, 2011).

En entornos de programación por bloques las instrucciones están representadas por bloques. Un programa se construye arrastrando y encastrando bloques en el orden apropiado para un determinado fin. En este modo de programar los conceptos abstractos tienen una representación visual. Además se tiene la ventaja de que el programa escrito está libre de errores de sintaxis tan frecuentes en otros lenguajes de programación. Esto permite a los usuarios focalizarse en la creación del programa. Es por ello que las herramientas de programación por bloques son cada vez más utilizadas en los cursos de iniciación a la programación.

En nuestro país, desde el año 2013 la

Fundación Sadosky lleva adelante la iniciativa Program.AR que promueve la inclusión del aprendizaje significativo de Ciencias de la Computación las escuelas de todo el territorio nacional. Para ello lleva adelante entre otras actividades, cursos de capacitación docente y desarrollo de material y recursos didácticos (Fundación Sadosky, 2013). En tal contexto, han desarrollado la aplicación *Pilas Bloques* que provee un entorno de programación por bloques. En *Pilas Bloques* se incluyen más de 40 actividades (desafíos) que abordan los principales conceptos de programación con niveles crecientes de dificultad.

Los desafíos incluidos en *Pilas Bloques* han sido diseñados para adoptar una metodología de enseñanza basada en problemas (Torp & Sage, 1998) que aplique una didáctica de la programación por indagación (Sanzo et. al, 2017). Para esta estrategia de enseñanza la indagación autodidacta es fundamental un rol activo del alumno, siendo el docente el que guía y asiste el proceso de aprendizaje del alumno.

En la carrera Ingeniería en Informática que se dicta en la Facultad de Tecnología y Ciencias Aplicadas de la Universidad nacional de Catamarca, la materia Programación I es la que mayor cantidad de cursadas requiere para alcanzar la condición de alumno regular. Lo cual constituye una evidencia de la dificultad para desarrollar las capacidades básicas asociadas al Pensamiento Computacional.

Problemática similar a la aquí planteada ha sido consignada por docentes investigadores de Universidades tanto de nuestro país como del exterior. Tal el caso de la Universidad Nacional de San Luis (Zúñiga, 2014), y de la Escuela de Ingeniería Informática de la chilena Universidad de Valparaíso (Muñoz, 2015). En ambos trabajos se expone a las herramientas de programación por bloques como un recurso útil para enseñanza en los primeros cursos de carreras de grado en Informática.

Líneas de Investigación, Desarrollo e Innovación

La línea de investigación aquí presentada trata de incorporar un novedoso enfoque para la enseñanza de nociones básicas de programación a estudiantes de Ingeniería en Informática.

El trabajo de investigación comprende el estudio de las herramientas y actividades apropiadas para introducir a conceptos de programación, el diseño de los instrumentos de recolección de datos para luego que, luego de realizada la intervención didáctica, hacer una evaluación del impacto de la innovación en la enseñanza a estudiantes de los primeros años de la carrera Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas de la U.N.Ca.

Resultados y Objetivos

Objetivo General

Promover el desarrollo de habilidades cognitivas básicas y específicas de pensamiento computacional y de la programación a alumnos de las cátedras Fundamentos de Informática (1º año) y Programación I (2º año) de la carrera Ingeniería en Informática de la Facultad de Tecnología y Ciencias Aplicadas (FTyCa) de la Universidad Nacional de Catamarca (UNCa).

Objetivos Específicos

- Facilitar el abordaje de conceptos relacionados con la Programación.
- Impulsar el desarrollo de habilidades del Pensamiento Computacional.
- Evaluar el nivel mejora en lo cognitivo, en lo actitudinal y motivacional de los alumnos en base a la intervención didáctica propuesta.

Resultados esperados

Mediante el presente trabajo de investigación se busca la formación y consolidación de un equipo de docentes y alumnos de la Facultad de Tecnología y Ciencias Aplicadas que se especialice en la enseñanza de los fundamentos de la programación a través de una didáctica por indagación utilizando plataformas educativas de programación por bloques. A partir de la experiencia a desarrollar se pretende continuar en años sucesivos con este tipo de actividades con nuevos grupos de estudiantes en las cátedras mencionadas.

Se espera observar que a través de la pauta pedagógica y recursos didácticos utilizados, los alumnos logren aprendizajes significativos de los conceptos elementales de programación fortaleciendo las capacidades relacionadas a las 3 dimensiones del Pensamiento Computacional (conceptos, prácticas y perspectivas computacionales) de modo pueda incorporarlas de manera natural y hacer uso de estas capacidades en todos los aspectos de su formación profesional.

Formación de Recursos Humanos

Los objetivos y actividades del presente proyecto, llevarán a que se realice capacitación permanente de todos los integrantes del Proyecto de Investigación en temáticas de Pensamiento Computacional y Didáctica de la Programación, aportando a la formación complementaria en las carreras de posgrado que cursan algunos integrantes del proyecto:

- Maestría en Informática de la Universidad de Santo Tomás de Aquino (Tucumán) - Lic. Valeria Póliche,
- Maestría en Ingeniería de Software. - Lic. Oscar Quinteros bajo la co-dirección del Dr. Hernán Ahumada,
- Carrera de Especialización en Docencia Universitaria de Disciplinas Tecnológicas de la Facultad de Ciencias Agrarias de la Universidad Nacional de Catamarca - Lic. Daniel Rivas.
- Incorporación de cuatro docentes no categorizados a los efectos que se inicien en Proyectos de Investigación acreditados: Lic. Oscar Quinteros, Esp. Lic Nelson Contreras, Esp. Lic. Marta Miranda e Ing. Javier Favore.

Los integrantes de este Proyecto tienen previsto participar de eventos científicos con el doble propósito de capacitarse en temáticas específicas concernientes al proyecto y para dar a conocer los resultados de las investigaciones realizadas. Eventos Científicos Nacionales a participar: WICC (Workshop de Investigadores en Ciencias de la Computación), Congreso Argentino de Tecnología en Educación y Educación en Tecnología (TE&ET), Simposio Argentino de Educación en Informática (SAEI), Jornadas Argentinas de Didáctica de la Programación (JADIPRO).

Referencias

ISTE and CSTA (2011). Computer Science Teachers Association and the International Society for Technology in Education. “Pensamiento Computacional, Caja de Herramientas”. Eduteka. Disponible en: http://www.eduteka.org/modulos/9/272/2062/1.

Brennan, K., & Resnick, M. (2012). New frameworks for studying and assessing the development of computational thinking. In Proceedings of the 2012 annual meeting of the American Educational Research Association, Vancouver, Canada (pp. 1-25).

Dann, W., Cooper S. & Pausch, R. (2011). Learning to Program with Alice. Prentice Hall

Fundación Sadosky (2012). "Informe CC 2016. Una propuesta para refundar la enseñanza de la computación en las escuelas argentinas", Buenos Aires, http://www.fundacionsadosky.org.ar/wp-content/uploads/2014/06/cc-2016.pdf.

- Levis, D. (2007). Enseñar y Aprender con informática/ Enseñar y aprender informática. Medios Informáticos en la escuela Argentina. In R. Cabello & D. Levis (Eds.), Medios Informáticos en la Educación a principios del siglo XXI (pp. 21–50). Buenos Aires-Argentina: Prometeo. -Martínez López, P. E. (2013). Las bases conceptuales de la programación. Una nueva forma de aprender a programar. Creative Commons.

Martínez López, P. E., Bonelli, E. A. & Sawady O'Connor, F.A. (2012). El nombre verdadero de la programación. Una concepción de la enseñanza de la programación para la sociedad de la información. Anales del 10mo Simposio de la Sociedad de la Información (SSI'12), dentro de las 41ras Jornadas Argentinas de Informática (JAIIO '12), 1–23. ISSN 1850-2830.

Muñoz, R., Barcelos, T. S., Villarroel, R., Barría, M., Becerra, C., Noel, R., & Frango Silveira, I. (2015). Uso de Scratch y Lego Mindstorms como apoyo a la docencia en Fundamentos de programación. In Actas de las XXI Jornadas de la Enseñanza Universitaria de la Informática (pp. 248-254). Universitat Oberta La Salle.

Sawady O’Connor, F.A. & Factorovich, P. (2013). Cuaderno Teórico de Scratch. Fundación Sadosky

Torp, L., & Sage, S. (1998). El Aprendizaje Basado en Problemas. (E. Litwin, Ed.). Buenos Aires: Amorrortu.

Wing, J.M. (2006). Computational thinking. Communications of the ACM, 49(3), 33–35.

Zúñiga, M. E., Rosas, M. V., Fernández, J., & Guerrero, R. A. (2014). El desarrollo del pensamiento computacional para la resolución de problemas en la enseñanza inicial de la programación. In XVI Workshop de Investigadores en Ciencias de la Computación

PROPUESTA DIDÁCTICA INHERENTE AL ÁREA DE CIENCIA DE DATOS

Mag. Raúl Oscar Klenzi, Mag. María Alejandra Malberti, Mag. Graciela Elida Beguerí

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
e-mail: {rauloscarklenzi,amalberti,grabeda}@gmail.com

RESUMEN

En el presente trabajo se expone una propuesta de cómo introducir contenidos inherentes al área de Ciencia de Datos en el marco de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información pertenecientes al Departamento de Informática –DI– de la Facultad de Ciencias Exactas Físicas y Naturales de la Universidad Nacional de San Juan –FCEFN–UNSJ–. Se plantea como una experiencia didáctica y surge de la revisión curricular de las carreras de grado. KNIME Analytics es el software de aprendizaje de máquina que utiliza tanto el grupo de investigación como la asignatura Inteligencia Artificial –IA–, marco en la que se realiza la experiencia. La asignatura IA pertenece a ambas carreras. En este caso se expone cómo evaluar distintos tipos de datos, relevados u obtenidos por los alumnos, para la finalidad planteada.

Palabras clave: Ciencia de Datos, Ingeniería de Datos, KNIME Analytics, Software Libre, Visualización, Educación Superior.

CONTEXTO

La temática está inserta en el proyecto "Visualización y Deep Learning en Ciencia de Datos" el cual se enmarca en el Laboratorio de Sistemas Inteligentes para Extracción de Conocimiento en Datos Masivos del Instituto de Informática de la Facultad de Exactas Físicas y Naturales de la Universidad Nacional de San Juan –FCEFN–UNSJ–.

Las líneas de investigación se presentan como una continuidad de los proyectos "Minería de Datos en la Determinación de Patrones de Uso y Perfiles de Usuario" y "Búsqueda de Conocimientos en Datos Masivos" aprobados y subsidiados por CICITCA.

1. INTRODUCCIÓN

La ciencia de datos y la ingeniería de datos son dos ramas diferentes dentro del paradigma de big data, un enfoque en el que se capturan, procesan a velocidades enormes, variedad y volúmenes de datos estructurados, no estructurados y semiestructurados, utilizando un conjunto de técnicas y tecnologías completamente novedosas en comparación con las que se utilizaron en décadas anteriores. Ambas son útiles para derivar conocimiento a partir de datos en bruto. Son elementos esenciales para cualquier sistema integral de apoyo a la toma de decisiones y extremadamente importantes a la hora de formular estrategias sólidas para la gestión empresarial.

Ciencia de Datos. La ciencia de datos se puede pensar como el dominio científico dedicado al descubrimiento de conocimiento mediante análisis de datos. El término específico de dominio, se refiere al sector de

la industria o dominio temático que los métodos de ciencia de datos utilizan para explorar. Los científicos de datos aplican técnicas matemáticas y enfoques algorítmicos para derivar soluciones a problemas empresariales y científicos complejos. Tanto en los negocios como en la ciencia, los métodos de ciencia de datos pueden proporcionar capacidades de toma de decisiones más robustas.

Ingeniería de datos. Es el dominio de ingeniería dedicado a superar los cuellos de botella de procesamiento de datos y problemas de manejo de datos para aplicaciones que utilizan Big Data. Los ingenieros de datos emplean la informática e ingeniería de software para modelar sistemas y resolver problemas con el manejo y la manipulación de grandes conjuntos de datos. Para lo cual se nutre de entornos de procesamiento en tiempo real y plataformas de procesamiento paralelo masivo (MPP), así como de sistemas de administración de bases de datos relacionales. En términos simples, con respecto a la ciencia de datos, el propósito de la ingeniería de datos es diseñar soluciones de big data mediante el desarrollo de plataformas de procesamiento de datos coherentes, modulares y escalables a partir de las cuales los científicos de datos puedan obtener posteriormente información valiosa (Pierson, 2015).

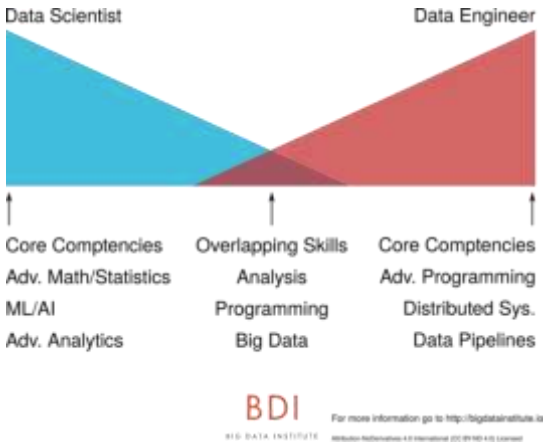


Figura 1. Competencias básicas de los científicos de datos e ingenieros de datos y sus habilidades superpuestas. https://www.oreilly.com

La relación entre un ingeniero de datos y un científico de datos, se observa en la figura 1.

Jesse Anderson, si bien explica también la diferencia entre un científico de datos con un ingeniero de datos, menciona que el profesional que está capacitado para ser competente tanto en ingeniería de datos como en ciencia de datos es el ingeniero en aprendizaje automático, lo que muestra la figura 2.

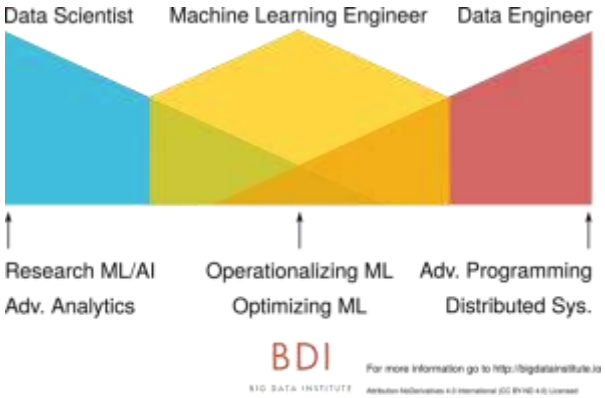


Figura 2. Diagrama que muestra dónde encaja un ingeniero de aprendizaje automático con un científico de datos e ingeniero de datos. https://www.oreilly.com

Si bien existen herramientas de software que pueden involucrar tanto las tareas de un ingeniero como de un científico de datos, el grupo de investigación viene utilizando KNIME Analytics, desde hace varios años, por ser una herramienta de software libre que, en forma modular permite incorporar todos los pasos involucrados en el proceso de ciencia de datos. Además esta plataforma basada en Eclipse, permite la integración tanto con proyectos de Eclipse como BIRT, DTP, etc., como con diferentes herramientas de recopilación, gestión y análisis de información. Proporciona un repositorio de módulos fáciles de usar y adicionalmente a las técnicas estándares de minería de datos, añade los algoritmos más actuales de análisis tales como los de deep learning (Salgueiro, 2016).

Con la intención de que los alumnos experimenten las actividades tanto de un ingeniero de datos como de un científico de

datos, es que en la asignatura IA se planificó y llevó adelante una actividad consistente en estudiar algunos casos provistos en el software y luego realizar una implementación con datos generados u obtenidos por ellos.

En la experiencia práctica se consideraron entre otros, datos georreferenciados, imágenes, texto extraído de redes sociales, interacción de KNIME con otros lenguajes de programación etc. Particularmente se mencionan los referentes a los datos georreferenciados e imágenes: german-bike-routes.gpx y 04_Car_Counting.knwf los que luego de ser estudiados fueron emulados con datos propios.

Es de destacar que para el primero de los ejemplos los alumnos procesaron datos georreferenciados inherentes a diferentes zonas de la provincia de San Juan, rescatados desde sitios de internet o registrados desde caminatas por medio de app's de dispositivos móviles, y luego graficaron los recorridos desde módulos específicos de KNIME.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Entre las líneas más importantes de investigación se mencionan:

- Ciencia de Datos, principalmente con Visualización de Información
- Deep Learning
- Herramientas de software libre para arquitecturas secuenciales, paralelas y distribuidas.

El proyecto tiene contemplado contribuir con el análisis y revisión curricular de las carreras de grado del DI con el fin de vincular las áreas disciplinares e incorporar, integrar y explotar simultáneamente distintas herramientas libres, provenientes de los aportes de las Tecnologías de la Información y la Comunicación, teniendo en cuenta aspectos de Interacción Humano Computador enriquecidos por teorías de la percepción humana.

3. RESULTADOS OBTENIDOS Y ESPERADOS

El proyecto se encuentra en su fase de desarrollo y al momento se está trabajando:

- Con datos provistos por una empresa de mantenimiento del alumbrado público de la ciudad de San Juan y se incursiona en aspectos relativos a geolocalización tales como análisis de coordenadas geográficas, su representación, manipulación e interpretación por parte de herramientas informáticas, así como su incorporación en un proceso de minería de datos (MD).
- En el área de Netmining (área de MD encargada de extraer conocimiento contextualizado a grafos mediante el cual se representa el contenido de redes sociales), se indagan datos provenientes de la Biblioteca de la FCEFN.
- Procesamiento de texto e imágenes mediante la utilización de algoritmos de deep learning.
- Determinación de diferentes métricas, desde series temporales, asociadas a entidades financieras y comerciales sobre datos extraídos de internet.

Como tareas de DS se reconocieron y georreferenciaron diferentes puntos de la ciudad de San Juan, producto de reclamos de usuarios de empresa de mantenimiento de alumbrado público. Con éstos, se emuló el problema del viajante de comercio y mediante Tabú Search fue resuelta la problemática, respetando los sentidos de circulación de cada trayecto en particular.

Como en cada iteración de la búsqueda tabú se debe acceder a los servicios de google a efectos de encontrar el camino mínimo y atento a resolver problemas de accesibilidad, se decidió utilizar los datos del proyecto Open Street Map –OSM-. Con los datos

correspondientes a la ciudad de San Juan - República Argentina, se implementó un servidor propio de mapas utilizando un contenedor (docker) y así tener un acceso permanente y rápido dentro de la intranet de la facultad. Esta alternativa contempla al DI.

Mientras se llevaron adelante las tareas planificadas en la cátedra, se observó un mayor interés por parte de los alumnos de participar en los diferentes proyectos de investigación enmarcados en el área temática de Ciencia de Datos y Big Data.

Así mismo, el equipo de la cátedra Inteligencia Artificial, mantiene hasta el presente esta modalidad como alternativa de enseñanza-aprendizaje referida a la extracción de conocimiento en datos

4. FORMACIÓN DE RECURSOS HUMANOS

Los integrantes del proyecto son docentes-investigadores que se desempeñan en las asignaturas Probabilidad y Estadística, Sistemas de Datos, Bases de Datos, Inteligencia Artificial y Programación Web, entre otras, de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información pertenecientes al -DI- de la -FCEFN-UNSJ-.

También se encuentran insertos en el proyecto alumnos, tesistas y becarios de las carreras del DI.

Los alumnos de grado se hallan realizando sus trabajos finales en las líneas de investigación mencionadas.

Miembros del equipo dirigen y asesoran trabajos finales de grado y posgrado; así como becas de Investigación, categoría Alumnos Avanzados CICITCA y CIN.

Es de destacar, en los casos detallados en el presente documento, la importante colaboración brindada por los alumnos Ana Paula Molina, Pablo Gómez, Sergio Quiroga y Juan Olivares, durante el desarrollo de la asignatura IA.

5. **BIBLIOGRAFÍA**

- Alcalde, I. (2015). *Visualización de la información: de los datos al conocimiento*. Editorial UOC.
- Anderson, Jesse (2018) Data engineers vs. data scientists. https://www.oreilly.com/ideas/data-engineers-vs-data-scientists
- Beguerí, G., & Malberti, A. (2017). Minería de Datos y una Aplicación en la Educación Superior. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*.
- Boulic, R., & Renault, O. (1991). 3d hierarchies for animation. *New Trends in Animation and Visualization, Edited by Nadia Magnenat-Thalmann and Daniel Thalmann, John Wiley & Sons ltd., England*.
- Cervone, G., Lin, J., &Waters, N. (Eds.). (2014). *Data Mining for Geoinformatics*. Springer New York.
- Karimi, H. A. (2014). *Big Data: techniques and technologies in geoinformatics*. Crc Press.
- Klenzi, R. O., Malberti, M. A., & Beguerí, G. E. (2018). Visualization in a Data Mining Environment from a Human Computer Interaction Perspective. *Computación y Sistemas*, *22*(1).
- KNIME (2016). KNIME Analytics versión 3.2.1. Software.
- Ortega, Matías."El negocio de datos en Argentina tiene un potencial de u$s 19.000 millones" https://www.ambito.com/el-negocio-datos-argentina-tiene-un-potencial-us-19000-millones-n4011082 Enero 2018

- Pierson, L. (2015). *Data sciencefordummies*. John Wiley&Sons.
- Salgueiro, A. P. (2016, September). Plataforma de Internet de las cosas para la enseñanza y la investigación. In *XIV Congreso Internacional de Información Info'2016*.
- Smith, A., & Jones, B. (1999). On the complexity of computing. *Advances in Computer Science*, 555-566.

Un marco para la enseñanza de ingeniería distribuida de requisitos en las universidades

Gustavo Sevilla[1], Sergio Zapata[1], Estela Torres[1], Fáber Giraldo[2], Facundo Gallardo[1]

[1]Instituto de Informática / Facultad de Ciencias Exactas, Físicas y Naturales / Universidad Nacional de San Juan; {gsevilla,szapata,etorres,fgallardo}@iinfo.unsj.edu.ar

[2]SINFOCI / Universidad de Quindío, Armenia, Colombia; fdgiraldo@uniquindio.edu.co

RESUMEN

La colaboración y las tecnologías de comunicación son de fundamental importancia para alcanzar con éxito las metas comunes que se propone un equipo de trabajo global o virtual de desarrollo de software. Los problemas de distancia física, horaria y cultural afectan fuertemente aspectos sociales del equipo global, como la cohesión grupal, la confianza interpersonal, el intercambio de conocimiento, la interdependencia laboral, etc. No se han producido avances significativos en propuestas que mitiguen estos problemas dentro de un proceso de desarrollo de software global o distribuido.

Por otra parte la mayoría de los modelos de enseñanza actuales de Ingeniería de requisitos de Software (IRS) en las universidades no consideran estos nuevos escenarios distribuidos.

El presente proyecto intenta hacer aportes para atender la falencia mencionada, presentando un marco para la enseñanza de IRS distribuida en las universidades.

Palabras clave: Tecnología Aplicada en Educación, Análisis de Requisitos, Ingeniería de Requisitos Distribuidos, Experiencias Prácticas.

CONTEXTO

El presente trabajo esta inserto en el proyecto de investigación titulado "Colaboración, Confianza y Desarrollo Global de Software" (Código 21/E1061), financiado por la Universidad Nacional de San Juan (UNSJ) y cuyo desarrollo se realiza en el Instituto de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales.

1. INTRODUCCION

El proceso de desarrollo de software tiende, con mayor énfasis, a distribuirse globalmente donde los protagonistas de este proceso (partes interesadas) están dispersos geográficamente. Esta forma moderna de trabajo se llama Desarrollo Global de Software (DGS). Existen varias razones para este escenario: reducción de costos, mejor uso global de la escasa disponibilidad de recursos humanos, más horas de trabajo disponibles en relación con las diferentes zonas horarias, incentivos para invertir en mercados emergentes y un crecimiento significativo en la demanda de software en todo el mundo [1]. Es aceptable sostener que estos nuevos contextos de trabajo implicarán adaptaciones en las formas y contenidos de la enseñanza de Ingeniería de Software, buscando preparar profesionales para un contexto laboral globalizado.

Se sabe que el éxito de un proyecto de software depende en gran medida de una ejecución correcta del proceso de IRS. Tener una definición correcta de los requisitos del software es clave para obtener un producto de calidad que cumpla con las necesidades y expectativas del cliente / usuario. La necesidad de tener un proceso de IRS dentro del proceso de Ingeniería de Software es imperativa para obtener productos de calidad [2]. La IRS es fundamentalmente un proceso de comunicación entre un especialista en requisitos y el cliente [3].

La etapa de elicitación de requisitos es la más crítica de todas las fases en el desarrollo de software, ya que los errores cometidos en esta etapa son más costosos y difíciles de resolver debido a su impacto en las otras etapas [4], [5]. Asimismo, la calidad de los requisitos mejora en la medida en que hay una mayor participación de los interesados. Esta participación se dificulta cuando el proceso de IRS se lleva a cabo en un entorno distribuido de desarrollo, es decir, cuando los actores están geográficamente dispersos. De acuerdo con el trabajo de Damian [6], lograr la participación adecuada de los usuarios del sistema y el personal de campo es uno de los desafíos identificados en estos escenarios y se debe tanto a la comunicación inadecuada como a la diferencia horaria.

Lo expuesto justifica la necesidad de prestar atención a la enseñanza de la IRS en entornos de DGS. El modelo de enseñanza debe promover y facilitar la colaboración de los involucrados, siendo las técnicas de elicitación de requisitos y las herramientas de comunicación elementos clave para cumplir con este propósito.

La mayoría de los modelos de enseñanza actuales de IRS no consideran estos nuevos escenarios distribuidos. Además, rara vez alcanzan experiencia práctica profesional, y la enseñanza se centra frecuentemente en la teoría, mientras que los estudiantes rara vez se involucran en proyectos reales [7].

Este trabajo presenta un marco para la enseñanza de IRS distribuida. El marco aborda la ejecución de prácticas de IRS con equipos formados por estudiantes y profesores de universidades latinoamericanas. El modelo está respaldado por estrategias de colaboración basadas en Wiki, videoconferencia (VC) y correo electrónico. El modelo se aplica a las prácticas de elicitación de requisitos de software e incluye actividades de comunicación e interacción entre todos los participantes distribuidos de los proyectos de software (es decir, ingenieros de requisitos, desarrolladores, usuarios, etc.).

2. LINEAS DE INVESTIGACION Y DESARROLLO

La línea de investigación sobre la cual se ejecuta el presente trabajo intenta generar conocimiento científico sobre los aspectos sociales involucrados en equipos distribuidos o globales de desarrollo de software. Seaman [8] argumenta que la Ingeniería de Software es una compleja disciplina pues combina tanto aspectos técnicos como aspectos humanos.

Cada vez es más reconocida la importancia que tienen los vínculos sociales de los miembros de los equipos de desarrollo y su impacto en la efectividad del proceso de software. La complejidad de entender estas relaciones humanas se incrementa cuando los miembros de esos equipos están distribuidos y las distancias físicas, temporales, culturales, de lenguaje afectan las comunicaciones e interacciones de los mismos. Bajo este contexto de investigación es que desenvuelven las actividades del presente trabajo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Al momento de llevar a cabo los trabajos experimentales descritos en esta sección, la enseñanza de la IRS en las carreras de Sistemas de la mayoría de las universidades latinoamericanas estaba enmarcada en escenarios colocalizados. Teniendo en cuenta esta situación descrita y la aparición de los nuevos escenarios de DGS, los experimentos consideraron la necesidad de investigar las técnicas de elicitación de requisitos y las herramientas de comunicación adecuadas para el proceso de ingeniería de requisitos en estos escenarios.

En el caso del primer experimento [9], se pretendía adquirir conocimientos de la etapa de elicitación distribuida de los requisitos de software en los nuevos escenarios de DGS, determinando las técnicas más convenientes de elicitación a utilizar. Se realizó un experimento controlado con dos factores: contexto de elicitación (distribuido/colocalizado) y técnicas de elicitación utilizadas (tres combinaciones diferentes de técnicas). Así, se

ejecutaron dos fases experimentales similares. La primera se aplicó a un contexto de desarrollo distribuido de software, mientras que la segunda se aplicó en un contexto colocalizado tradicional. El diseño experimental fue finalmente un diseño factorial 2x3.

Los resultados promedio del Nivel de Efectividad QRSD de cada combinación de técnica de elicitación, se pueden ver en las Tablas 1 y 2, agrupándolos por las técnicas de elicitación aplicadas en los dos procesos de elicitación realizados, en otras palabras, distribuidos y colocalizados, respectivamente.

Tabla 1. Promedio de resultados de la fase distribuida

Técnica	QRSD Promedio
Entrevista	66.76
Entrevista + Cuestionario	74.94
Entrevista + Brainstorming	68.94

Tabla 2. Promedio de resultados de la fase colocalizada

Técnica	QRSD Promedio
Entrevista	77.12
Entrevista + Cuestionario	77.54
Entrevista + Brainstorming	77.44

De los datos experimentales obtenidos surgue que la combinación de técnicas de elicitación de requisitos más efectivas en un contexto distribuido sería la Entrevista-Cuestionario, que alcanza un nivel de efectividad que supera en el 9% la siguiente combinación de técnicas más efectiva, que fue la Entrevista-Brainstorming.

En el segundo experimento [10] alumnos avanzados hacían de Ingenieros de Requisitos e interactuaban con profesores que hacían de Clientes para obtener el Documento de Requisitos de Software (DRS) utilizando una Wiki en el proceso IRS en un escenario de desarrollo distribuido de software.

La Tabla 3 muestra los resultados obtenidos por cada grupo de IRS con respecto a tres atributos de calidad del Documento de requisitos de software: Completitud, Ausencia de ambigüedad y Precisión. Otro índice medido es el grado o nivel de uso de la herramienta wiki por cada grupo.

Tabla 3. Resultados Obtenidos

Grupo	Índice Uso de Wiki	Índice Completitud q1	Índice No-Ambiguo q2	Índice Precisión q3
G1	1.725	0.500	0.800	0.650
G2	1.529	0.540	0.941	0.765
G3	2.345	0.520	0.840	0.800
G4	2.300	0.460	0.700	0.750
G5	2.020	0.600	0.800	0.520
G6	1.618	0.500	0.618	0.382
G7	1.750	0.780	1.000	0.538
G8	1.339	0.520	0.731	0.346
G9	1.283	0.520	0.652	0.174

Sobre la base de la evidencia experimental obtenida y presentada anteriormente, respecto a las técnicas de elicitación de requisitos y las herramientas de comunicación utilizadas en los escenarios DGS, se propone un marco para la enseñanza de IRS en las universidades para escenarios de DGS donde el cliente/usuario está ubicado de forma remota con respecto al ingeniero de requisitos.

El marco propone, ver Figura 1, en base a las experiencias realizadas, el uso de técnicas de elicitación de requisitos y herramientas de comunicación en diferentes fases del proceso IRS. La propuesta trata de mantener una interacción continua con el cliente/usuario durante todo el proceso, teniendo el mismo diferentes grados de protagonismo en las diferentes fases. En este sentido, la wiki, específicamente configurada para el proceso IRS, es una herramienta de comunicación y repositorio que registra el progreso del proceso. Específicamente, los requisitos de software establecidos con sus comentarios asociados.

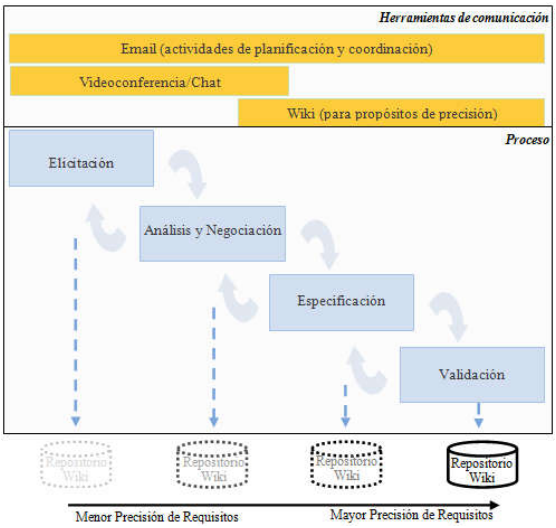


Figura 1: Marco propuesto

La Wiki es accedida tanto por el ingeniero de requisitos como por el cliente/usuario. La información que se genera en cada fase del proceso se almacena en la wiki que soporta el modelo de proceso.

Sobre la base de la evidencia experimental presentada en este documento, se propone que la comunicación asíncrona proporcionada por el correo electrónico se utilice durante todo el proceso de la IRS, en todas las fases. Normalmente, los correos electrónicos apoyan el intercambio de mensajes de planificación y coordinación entre las partes del proceso. Por ejemplo, solicitudes de VC, notificaciones de finalización de actividades, alarmas de tareas vencidas y difusión de aspectos relacionados con el progreso del proyecto, etc. Es por esta razón que el uso del correo electrónico aparece a lo largo de todo el proceso.

Al final del proceso, el conjunto de requisitos de software que formará el núcleo DRS del sistema permanecerá registrado en el Repositorio Wiki.

En el marco propuesto, el proceso de IRS se divide en cuatro fases diferentes:

-Fase de Elicitación: El marco establece que en esta primera fase del proceso, la técnica de la entrevista debe realizarse mediante VC. Este primer contacto entre el cliente/usuario y el ingeniero de requisitos apunta a obtener una definición general, no precisa, del problema a resolver. Esta es la razón por la cual, como muestran los experimentos, las entrevistas no estructuradas son adecuadas para esta fase. Con este fin, la VC es una herramienta de comunicación de una riqueza comunicativa relativamente alta que soporta adecuadamente una entrevista remota. A medida que el conocimiento del problema progresa, se puede hacer el uso de cuestionarios para obtener una mayor precisión. Se pueden administrar de manera adecuada a través de una herramienta de comunicación asíncrona como el correo electrónico o de herramientas web de uso específico para cuestionarios como los formularios de Google.

-Fase de Análisis y Negociación: el marco establece que esta fase del proceso tiene como objetivo analizar la claridad, pertinencia y viabilidad técnica de los requisitos del software. Para este propósito, es necesario interactuar con el cliente/usuario para negociar aspectos relacionados con el esfuerzo y la viabilidad técnica de la implementación de los requisitos. En esta fase, las entrevistas deben realizarse a través de VCs. Como en todas las fases, los avances en términos de definición de requisitos de software se registran en el Repositorio Wiki. Al final de esta segunda fase es cuando la herramienta wiki como un medio de comunicación asíncrona entre el cliente/usuario y el ingeniero de requisitos comienza a tomar preponderancia.

-Fase de Especificación: esta fase del proceso tiene como objetivo establecer un conjunto de requisitos de software bien formulados, es decir, requisitos con un atributo de calidad suficiente para garantizar el éxito de las últimas etapas del proceso de desarrollo del software. El marco propone el uso de la herramienta wiki en esta fase, así el ingeniero de requisitos puede revisar y editar los requisitos para mejorar su calidad, mientras que el cliente/usuario puede verificar, modificar y validar los cambios. En esta fase, comienza un tipo de interacción asíncrona, textual, reflexiva y focal entre las partes, por lo que las VC ya no son útiles, convirtiendo a la wiki en la herramienta adecuada. Esta fase requiere un trabajo más detallado sobre los

requisitos, especialmente para darles precisión. Aquí es donde la herramienta wiki brinda un gran beneficio como lo demuestran los datos experimentales mencionados en este documento.

-Fase de Validación: Finalmente, en esta fase, ambas partes del proceso, cliente/usuario e ingeniero de requisitos, validan el conjunto de requisitos de software identificados y expuestos claramente en el Repositorio Wiki. Nuevamente, esta herramienta soporta estas tareas de revisión integral y validación.

La motivación adicional de los estudiantes de participar en estas prácticas junto con compañeros de otros países es un factor adicional significativo que contribuye positivamente al proceso de aprendizaje. El uso de herramientas de comunicación estándar, fácilmente disponibles en las universidades latinoamericanas, hace que el marco propuesto sea factible de implementar.

Resta aún realizar validaciones para efectuar afirmaciones sobre la utilidad del marco.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo es integrado por: Un Profesor Adjunto, un Profesor Titular, una Profesora JTP y un Ayudante alumno.

Esta línea de investigación permitió en el año 2018 obtener el título de Magister en Informática a uno de los autores de este artículo y a su vez constituye uno de los pilares para su estudio de doctorado próximo y una tesina de grado en desarrollo.

5. BIBLIOGRAFIA

[1] J. D. Herbsleb and D. Moitra, “Guest editors’ introduction: Global software development,” in IEEE Softw. 18, 2 (Mar. 2001), 16–20.

[2] A. Duran Toro, Un entorno metodológico de ingeniería de requisitos para sistemas de información, 2000.

[3] A. Abran, J. W. Moore, P. Bourque, R. Dupuis and L. L. Tripp, Guide to the Software Engineering Body of Knowledge (SWEBOK) version 3 public review. IEEE, 2013. ISO Technical Report ISO/IEC TR 19759.

[4] R. W. Ferguson and G. Lami, “An empirical study on the relationship between defective requirements and test failures,” in Software Engineering Workshop, 2006. SEW’06. 30th Annual IEEE/NASA (2006), IEEE, pp. 7–10.

[5] P. Rajagopal, R. Lee, T. Ahlswede, C. C. Chiang and D. Karolak, “A new approach for software requirements elicitation,” in Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing, 2005 and First ACIS International Workshop on Self-Assembling Wireless Networks. SNPD/SAWN 2005. Sixth International Conference on (2005), IEEE, pp. 32–42.

[6] D. Damian & D. Zowghi, “RE challenges in multi-site software development organisations,” in Requirements Engineering 8.3, 2003, 149-160.

[7] U. Bellur, “An academic perspective on globalization in the software industry,” in Computer Software and Applications Conference, 2006, COMPSAC’06. 30th Annual International (2006), vol. 1, IEEE, pp. 53–54.

[8] C. B. Seaman, "Qualitative Methods". In: SHULL et al. (eds.), Guide to Advanced Empirical Software Engineering, Chapter 2, Springer, 2008.

[9] S. Zapata, E. Torres, G. Sevilla, L. Aballay and M. Reus, “Effectiveness of traditional software requirement elicitation techniques applied in distributed software development scenarios,” in CLEI 2012, IEEE, pp. 1–7.

[10] G. Sevilla, S. Zapata, E. Torres and C. A. Collazos, “Using wikis as collaborative strategy to support software requirements elicitation,” in Computing Colombian Conference (9CCC), 2014 9th (2014), IEEE, pp. 54–61.

Innovación en

Sistemas de Software

Agentes Inteligentes y Web Semántica

Christian Gimenez[1] Germán Braun[1,2] Angela Oyarzun[1]
Gaston Michelan[1] Emiliano Rios Gavagnin[1] Giuliano Marinelli[1]
Sandra Roger[1] Mario Moya[1] Claudio Vaucheret[1] Laura Cecchi[1]

email: {christian.gimenez,german.braun, gaston.michelan,giuliano.marinelli, roger, mario.moya, cv, lcecchi}@fi.uncoma.edu.ar
{angela.oyarzun, emiliano.rios}@est.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 Nacional de Investigaciones Científicas y Técnicas (CONICET)*

Resumen

En este trabajo se presentan las líneas de investigación desarrolladas en el Proyecto de Investigación Agentes Inteligentes y Web Semántica, de la Facultad de Informática, UNCo. En el proyecto, 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, para entre otros, dar soporte a comunidades de desarrollo de ontologías.
En este sentido, se ha desarrollado una herramientas Web denominada *crowd* que da soporte a la ingeniería ontológica asistida con razonamiento automático y que está siendo extendida con nuevas funcionalidades estudiadas en las diferentes líneas de investigación.

Palabras Clave: Ingeniería de Software basada en Conocimiento, Lógicas Descriptivas, Ontologías. Datos Enlazados. OVM

Contexto

Este trabajo está parcialmente financiado por la UNCo, 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 y de una Beca de Iniciación en la Investigación para Estudiantes. Asimismo, lo financia parcialmente el Consejo Interuniversitario Nacional (CIN) con dos Becas de Estímulo a las Vocaciones Científicas 2018 y 2019 y por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), en el contexto de una beca interna doctoral. El proyecto de investigación tiene una duración de cuatro años y ha comenzado en 2017.El proyecto se desarrolla en forma colaborativa con docentes-investigadores de la UNS.

1. Introducción

Actualmente, es esperable que las aplicaciones accedan a grandes cantidades de datos que se modifican en tiempo real, de manera constante y que se presentan en diferentes formatos, especialmente disponible a través de la Web. Muchas de estas aplicaciones ocurren bajo restricciones de tiempo críticas y en intensa interacción con el usuario. En este contexto, la representación conceptual de los dominios para la generación y extracción de información y conocimiento, es central en la toma de decisiones. Bajo este aspecto, el objetivo general del proyecto de investigación es el estudio y desarrollo

de técnicas de Inteligencia Artificial para dotar de inteligencia, conocimiento e interoperabilidad semántica a agentes inmersos en ambientes complejos.

Asimismo, la Web es un ambiente ideal para estos agentes que, junto con la generación y representación de conocimiento, las ontologías y el razonamiento automático, son fundamentales en la evolución de dicha Web, denominada Web Semántica. La meta de la Web Semántica es crear una Web de conocimiento en la cual la semántica del contenido es explicitada, permitiendo novedosas aplicaciones que combinan datos de sitios heterogéneos para, entre otros objetivos, mejorar la experiencia de los usuarios de acuerdo a sus necesidades. Esta red de funcionalidades basadas en conocimiento hará a dicho conocimiento procesable por una computadora, para soportar conductas inteligentes por parte de dichas máquinas.

Así en el proyecto, 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, para entre otros, dar soporte a comunidades de desarrollo de ontologías.

En esta dirección, se ha desarrollado *crowd*[1, 2], una herramienta Web, creada en respuesta a la complejidad inherente al modelado conceptual y ontológico.

En este trabajo se presentan las líneas de investigación desarrolladas en el Proyecto de Investigación Agentes Inteligentes y Web Semántica, de la Facultad de Informática. Se describen los resultados parciales alcanzados, avalados con publicaciones y los trabajos en desarrollo.

Agentes Inteligentes y Web Semántica tiene asociado el Proyecto de Extensión *Agentes Robots: Divulgando Computación en la Escuela Media*, cuyo objetivo es divulgar temas de la disciplina y enseñar a programar utilizando como eje motivador robots físicos, a través de la *Copa de Fútbol FAI: la definición es por penales*[3].

2. Línea de Investigación y Desarrollo

Motivados por el objetivo del proyecto, se delinearon varias líneas de investigación que confluyen explorando, entre otros, temas afines a la representación del conocimiento, las Lógicas Descriptivas (DLs), las Ontologías, la Ingeniería de Software basada en Conocimiento, la Ingeniería de Conocimiento y los Datos Enlazados y Abiertos. A continuación se las describen brevemente.

***(L1)* Formalización de Sistemas de Visualización Ontológicos con soporte de Razonamiento:** Durante varias décadas, un esfuerzo considerable ha sido invertido en ampliar el horizonte teórico de las DLs y los lenguajes de ontologías. Sin embargo, se ha relegado a un segundo plano el desarrollo de ambientes de ingeniería ontológica para simplificar la formalización y la validación de las propiedades de los modelos conceptuales sobre dominios de aplicación.

En esta línea se trabaja en la integración de lenguajes de modelado conceptual visuales con razonamiento basado en lógica y su utilidad en cuanto a tareas de ingeniería ontológica automáticas, y la validación de modelos. En este contexto, se formalizó un sistema teórico para la manipulación de modelos visuales, dando lugar a la definición de ontologías gráficas y se implementó *crowd* como instancia de esta formalización.

***(L2)* Lenguaje Gráfico para Consultar Ontologías:** En esta línea de investigación se propone desarrollar un lenguaje gráfico que represente las consultas SPARQL-DL a realizarse sobre una ontología. 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).

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. Las respuestas obtenidas, son mostradas al usuario en forma textual, pudiendo ser interpretadas gráficamente reemplazando las variables por los nombres de clases encontrados.

Para implementar este lenguaje, se continuará con el desarrollo de la herramienta Web denominada *crowd*, la que asiste al modelador en el diseño y la visualización de ontologías, actualmente a través del lenguaje de modelado conceptual UML y que soporta razonamiento automático para validar los modelos.

***(L3)* Alineamiento de Ontologías en un Ambiente Gráfico Web de Ingeniería Ontológica:** En la actualidad, los datos presentes en sistemas informáticos son muy abundante y están almacenados en distintos formatos. Las ontologías son indispensables en este escenario, ya que permiten una descripción formal de los datos y de su dominio. Por esta razón, es importante trabajar sobre la heterogeneidad presente en los modelos de información mediante el alineamiento de ontologías, el cual tiene como objetivo encontrar correspondencias entre entidades de modelos diferentes y, así, reducir la brecha semántica entre representaciones de un mismo dominio.

Esta línea de investigación propone trabajar sobre las técnicas de alineamiento de ontologías para integrarlas en la herramienta *crowd*. Se pretende desarrollar un diseño que permita la integración de ontologías diferentes en un entorno gráfico, mediante la identificación de correspondencias sintácticas y semánticas entre sus entidades.

***(L4)* Verbalización multilingüe del metamodelo UML:** Esta línea de investigación estudia la verbalización multilingüe del metamodelo UML, orientado hacia un subconjunto de primitivas de los diagramas de clases. Se utiliza una representación intermedia de los diagramas de clases en sentencias concretas de lógica de primer orden (FOL) para favorecer las verbalizaciones en diversos idiomas.

La implementación de esta formalización se realiza sobre *crowd* mediante una extensión que permita la incorporación de esta nueva funcionalidad. 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.

Un aspecto relevante, es que la traducción a lenguaje natural está destinada a ser multilingüe, a diferencia de la mayoría de las herramientas existentes, donde el lenguaje natural de salida es generalmente el inglés.

***(L5)* Algortimos de Layout para una Herramienta Web de Modelado Ontológico:** 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 visualización del contenido y facilite dicha tarea.

Debido a la complejidad inherente de modelos y ontologías de dominio en cuanto al tamaño de los diagramas y a los requerimientos de edición de los modeladores, algoritmos efectivos y eficientes de layout deben ser desarrollados para garantizar la legibilidad de los modelos. El objetivo de esta línea de investigación es investigar y diseñar algoritmos de layout. Para esto se profundizará en técnicas de visualización basadas en la teoría de *Crossing Number* y sus variantes, aplicado a los lenguajes clásicos de modelado conceptual como UML, EER y ORM.

Utilizaremos la herramienta *crowd* para la implementación de nuevos prototipos y simultáneamente este trabajo permitirá la ampliación de la herramienta.

***(L6)* Formalización de Modelos de Variabilidad Ortogonal:** En esta línea de investigación se trabaja en la formalización de Modelos de Variabilidad Ortogonal (OVM) en lenguajes formales como FOL y DL. Asimismo, tiene como objetivo el desarrollo de una herramienta Web, basada en un modelo cliente-servidor, que permita integrar el soporte gráfico para el diseño de OVM y el razonamiento automático

para validar dichos modelos. El objetivo de esta herramienta es solventar las limitaciones presentadas en las propuestas existentes basadas en razonadores SAT a través de la utilización de DLs. De igual manera, busca evitar la ocurrencia de errores en el diseño de los modelos al proveer una interfaz de usuario gráfica con un conjunto de primitivas del lenguaje OVM y las funcionalidades necesarias para la construcción y edición de los diagramas.

***(L7)* Datos Enlazados en el dominio de publicaciones científico:** Debido a la promulgación de las leyes en Argentina que exigen la publicación de datos abiertos, (Ley Nacional N° 26.899),existen sitios nacionales en los cuales se expone información en formatos de uso libre tales como JSON y .CSV. Teniendo en cuenta la importancia que comienza a tener el paradigma de Web Semántica en conjunto al concepto de Datos Enlazados, este trabajo propone el diseño de una herramienta que relacione mediante enlaces los datos ya existentes en esos formatos y permita enriquecer la información mediante la colaboración de diversas instituciones que unen sus bases de datos de información científica, manteniendo el estándar que se propone de enlaces y relaciones bajo el concepto de Web Semántica, promulgando la unicidad, consistencia y el libre acceso a la información.

3. Resultados Obtenidos y Trabajo Futuro

Actualmente, *crowd* permite crear modelos basados en un subconjunto de primitivas de UML. Soporta el razonamiento automático para determinar inconsistencias del modelo.Asimismo, es posible obtener la representación textual del modelo formalizado en OWL2 para ser utilizado con otras herramientas. Un prototipo está disponible en `crowd.fi.uncoma.edu.ar`.

Respecto de (L1), se está trabajando en la expansión de los conceptos teóricos para lograr mayor interacción gráfica y lógica y, también, en el desarrollo de nuevas funcionalidades de *crowd*: importación de ontologías OWL 2 y soporte a múltiple lenguajes gráficos de modelado conceptual. Asimismo, se ha comenzado con el desarrollo de la línea (L3) para abarcar más tareas asociadas a la ingeniería ontológica. En cuanto al layout automático (L5), se desarrolló un algoritmo genético [4] que permite disponer visualmente al grafo que representa al modelo con un crossing number adecuado. Actualmente, se está estudiando la adaptación de este algoritmo para que se comporte apropiadamente de acuerdo a los diferentes lenguajes de modelado conceptual soportados por *crowd*.

Definido el lenguaje visual propuesto en la línea (L2)[5], se está implementando una extensión de la arquitectura de *crowd*, que consiste en un diseño de interfaz para 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.

En cuanto a los resultados de la línea (L4), se diseñaron los esquemas de verbalización en lenguaje español, basados en $templates$[6]. Además, se integró al módulo de verbalización una herramienta que permite obtener la morfología gramatical (género y número) de cada uno de los nombres de clases y atributos. Asimismo, se estableció la conexión con un servicio web que realiza la conjugación verbal de los nombres de las asociaciones. Actualmente se está finalizando la fase de implementación, extendiendo a *crowd*. A futuro, se prevee la extensión a otros idiomas, mediante la creación de nuevos $templates$ de verbalización, y la correspondiente validación.

En esta línea de investigación (L6) se formalizó OVM utilizando lenguajes formales como FOL y DL[7]. Actualmente, se está desarrollando la herramienta Web, que permita integrar el soporte gráfico para el diseño de OVM y el razonamiento automático para validar dichos modelos[8].

Finalmente, en el marco de la línea (L7) se ha trabajado en procesar los datos existentes y convertirlos en un formato estándar (por convención RDF)[9]. En este sentido, es necesario además identificarlos unívocamente utilizando

URIs para mantener consistencia, reconciliarlos, enlazarlos y enriquecerlos con información relacionada mediante lenguajes de conocimiento como OWL, trabajo que se está desarrollando. Se plantea como trabajo futuro desarrollar una interfaz amigable para ser presentados y utilizados por los usuarios.

4. Formación de Recursos Humanos

Uno de los integrantes está finalizando el Doctorado en Ciencias de la Computación en la Universidad Nacional del Sur (beca interna doctoral CONICET). Otro de los miembros del proyecto, goza de una Becas de Investigación en la categorías de Graduado de Iniciación, UNCo y en el transcurso del 2019 comenzará a cursar un posgrado acreditado. Dos de los integrantes están cursando actualmente posgrados: Magíster en Ciencia, tecnología e Innovación de la UNRN y el Magíster en Ciencias de la Computación de la UNS.

Finalmente, los integrantes alumnos se encuentran terminando la carrera Licenciatura en Ciencias de la Computación. Tres de ellos están actualmente desarrollando su tesis. Uno de estos estudiantes goza de la Beca de Iniciación en la Investigación para Estudiantes de la UNCo y dos de ellos Becas de Estímulo a las Vocaciones Científicas 2018 y 2019 (CIN).

Referencias

[1] 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.

[2] 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.

[3] Jorge Rodriguez, Guillermo Grosso, Rafael Zurita, and Cecchi Laura. 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*, 2016.

[4] Giuliano Marinelli, Germán Braun, Laura Cecchi, and Pablo Rubén Fillottrani. Algoritmos de layout automático para una herramienta multivistas de modelado ontológico. In *XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste)*, 2018.

[5] Christian Gimenez, Germán Braun, Laura Cecchi, and Pablo Rubén Fillottrani. Towards a visual sparql-dl query builder. In *XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, Facultad de Ciencias Exactas - Universidad Nacional del Centro de la Provincia de Buenos Aires, 2018.

[6] Matías Garrido, Germán Braun, and Sandra Roger. Agentes inteligentes y web semántica: hacia la verbalización de un subconjunto de uml en una herramienta gráfica web. In *XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, ITBA, Buenos Aires)*, 2017.

[7] Germán Alejandro Braun, Matias Pol'la, Laura Andrea Cecchi, Agustina Buccella, Pablo R. Fillottrani, and Alejandra Cechich. A DL semantics for reasoning over ovm-based variability models. In *Proceedings of the 30th International Workshop on Description Logics, Montpellier, France, July 18-21, 2017.*, 2017.

[8] Angela Oyarzun, Germán Braun, Laura Cecchi, and Pablo Rubén Fillottrani. A graphical web tool with dl-based reasoning support over orthogonal variability models. In *XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, Facultad de Ciencias Exactas - Universidad Nacional del Centro de la Provincia de Buenos Aires, 2018.

[9] Gaston Michelan, Germán Braun, Laura Cecchi, and Pablo Rubén Fillottrani. Hacia la integración semántica de repositorios digitales abiertos en argentina. In *5to Congreso Nacional de Ingeniería Informática- Sistemas de Información(CONAIISI 2017)*, UTN - Facultad Regional Santa Fe, 2017.

ANDROID THINGS

Mg. Ing. Norma Beatriz Perez[(1)], Miguel Alfredo Bustos[(1)], Dr. Mario Marcelo Berón[(1)] & PhD. Pedro Rangel Henriques[(2)]
[(1)]Departamento de Informática – Facultad de Ciencias Físico Matemáticas y Naturales (FCFMyN) – Universidad Nacional de San Luis (UNSL)
Ejercito de los Andes 950, D5700HHW San Luis, +54-0266 4520300 – Int. 2103
[(2)] Departamento de Informática – Universidade do Minho – Braga, Portugal
{*mabustos, nbperez, mberon*}*@unsl.edu.ar - pedrorangelhenriques@gmail.com*

RESUMEN

En la actualidad, el *Internet of Things* (IoT) supera a los smartphones convirtiéndose en la categoría más importante de dispositivos conectados. Este ingente crecimiento de IoT impulsa a los desarrolladores de estas tecnologías evolucionar en el enfoque de la seguridad para proteger servicios vitales, información sensible de posibles robos, manipulación y pérdida de datos además de otros aspectos claves que permitan diseñar sistemas IoT confiables. Es por esto, que las organizaciones de desarrollo como Google, Apple se introducen en este mercado global incorporando a su staff plataformas emergentes de IoT. *Android Things* es una plataforma de vanguardia que provee Google para el desarrollo de dispositivos versátiles, de bajo costo, de calidad, entre otras características de relevancia que simplifican y satisfacen las necesidades de los usuarios que son cada vez más exigentes.

El presente trabajo describe una línea de investigación que estudia de manera exhaustiva la plataforma Android *Things* pensada principalmente para la implementación de objetos integrados e interconectados. Dicho estudio se fundamenta en la utilización de la plataforma Android *Mobile* como soporte para construir aplicaciones de Android *Things* confiable*s*. El análisis de estas tecnologías se basa detectar los principales lineamientos para proponer y desarrollar un metodología o herramienta que permita migrar aplicaciones existentes al nuevo mercado de Android *Things* de manera segura.

Palabras Claves: Android, Android *Things*, Aplicaciones*,* Seguridad.

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 de la Universidad Nacional de San Luis. Proyecto Nº P-031516 que es la continuación de diferentes proyectos de investigación, a través de los cuáles se ha logrado un importante vínculo con distintas universidades a nivel nacional e internacional. Además, se encuentra reconocido por el Programa de Incentivos. El segúndo 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 aprobado por el Ministerio de Ciencia, Tecnología e Innovación Productiva (MinCyT), y su código es PO/16/93.

1. INTRODUCCÓN

La aceptación global de la Internet; la gran dispersión de usuarios móviles; los miles de millones de dispositivos que se conectan entre sí con el auge de IoT; así como la enorme acumulación de información en la nube hacen que su uso e incorporación en los diferentes escenarios sea un reto primordial en los mercados emergentes donde se emplean estas tecnologías. El lanzamiento de la plataforma Android [1], [2] ha causado gran

impacto en el desarrollo de aplicaciones móviles [3] logrando una amplia aceptación en el mercado global donde cada vez las partes involucradas son más exigentes.

Hoy en día, la plataforma Android [4] se ha convertido en una alternativa dominante frente a otras plataformas. Esta plataforma se basa en el sistema operativo Linux por lo que es de código abierto y puede ser utilizado sin realizar costos adicionales a sus usuarios. Se destaca por ser adaptable a cualquier tipo de hardware como por ejemplo, Smartphone, Smartwatch, etc. y una amplia variedad de productos empotrados que utilizan este sistema operativo para llevar a cabo sus tareas. Las aplicaciones, en la plataforma Android, son desarrolladas en el lenguaje JAVA [5] esto permite que dichas aplicaciones pueden ser ejecutadas en cualquier dispositivo a través de la máquina virtual denominada Dalvik [6] ofreciendo portabilidad segura. Por otro lado, la arquitectura de Android [7] se basa en componentes "inspirados" en Internet. Por ejemplo, la interfaz de usuario es realizada en XML permitiendo que una misma aplicación se ejecute en diferentes pantallas bajo distintas dimensiones. Android incorpora servicios de localización [8], acceso a redes, bases de datos con SQL, síntesis de voz, multimedia, entre otros. La seguridad de Android [9] se provee a través de permisos que son otorgados por parte de sus usuarios.

La plataforma Android *Mobile* ofrece una manera intuitiva y novedosa de desarrollar e implementar potentes aplicaciones para diversos dispositivos. Android *Things* [10] extiende de la plataforma Android *Mobile* con el objetivo de ofrecer a los desarrolladores la posibilidad de construir objetos integrados e interconectados con características destacables como es la alta calidad, mayor seguridad, productos a escala, etc. Se destacan tres pilares primordiales de esta plataforma con respecto a Android *Mobile*: Arquitectura (incorporando una extensión del marco central con APIs adicionales que ofrece la biblioteca de soporte *Things*); Actualizaciones seguras (son administradas por la central de Google); Optimización de Sistema Operativo [11] (se dispone de una variante con el objetivo de ser utilizado en IoT [12]) y Hardware potente (accesible y de fácil integración). Es por esto, que la línea de investigación aborda el estudio de los pilares mencionados previamente.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la actualidad, para la implementación y desarrollo de aplicaciones IoT que utilizan como soporte la plataforma Android *Things* es una tarea intuitiva si el desarrollador a programado previamente bajo Android. Esto se debe a que la migración de aplicaciones desarrolladas en Android *Mobile* (que utilizan el entorno de desarrollo integrado Android Studio [13], [14], APIs, etc.) a Android Things resultan en una convergencia relativamente simple debido a que ambas plataformas comparten características similares por ser implementadas por la misma organización de desarrollo denominada Google. La popularidad de Android se debe a la amplia aceptación de sus usuarios en este mercado.

IoT está en pleno auge incorporando aplicaciones móviles que facilitan el uso, administración y control de los dispositivos que intervienen en un sistema IoT. Sin embargo, IoT trae asociado riesgos de seguridad, presentando en ocasiones fallas de carácter crítico pudiendo vulnerar a organizaciones, gobiernos, usuarios potenciales, entre otros.

Por otro lado, Android *Things* es una tecnología emergente que incorpora actualizaciones seguras con el objetivo de evitar perdida de información relevante, contrarrestar las amenazas, etc.

Lo mencionado en los párrafos anteriores, deja en evidencia que los desarrolladores de aplicaciones basado en

Android *Things* se enfrentan a un problema principal que es la creación de actualizaciones seguras y escalables ya que hacen uso de IoT.

Esta línea de investigación propone:

- Realizar un estudio profundo sobre los riesgos de seguridad que enfrentan los sistemas IoT.
- Realizar un estudio exhaustivo de Android *Things* incluyendo su arquitectura, SO, actualizaciones, etc.
- Integración de aplicaciones basadas en Android *Mobile* y Android *Things*.
- Determinar problemas de seguridad de estas tecnologías.

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 [15] para desarrollar mecanismos de seguridad en IoT se basa en tres áreas: (i) Confidencialidad de Datos (se utilizan mecanismos de cifrado de datos); (ii) Integridad de Datos (mecanismo cyclic redundancy check. Para evitar el acceso a datos sensibles [7]); disponibilidad de datos (métodos de copia de seguridad de redundancia y conmutación por error. Permiten acceso a sus recursos a quien corresponda en condiciones normales y adversas).

2.2 Android Things

La plataforma Android *Things* es un extensión de la plataforma Android *Mobile* por lo que tiene grandes similitudes en los componentes como por ejemplo: Android SDK, Android Studio, Play Services, FireBase, Google Cloud IoT Core favoreciendo a los desarrolladores de estas tecnologías su flexible y intuitiva convergencia. Se introducen, a continuación, los componentes principales de Android *Things*.

2.2.1 Arquitectura

La arquitectura de Android *Mobile* tradicional se muestra en la Figura 1. Donde se puede observar los núcleos y bibliotecas que son principalmente los encargados de habilitar el soporte del controlador de hardware.

Figura 1: Arquitectura de Android Mobile

Los frameworks de aplicaciones ofrecen un conjunto de APIs con finalidad de ser utilizadas por las mismas. Además, es importante mencionar que las aplicaciones proveen un uso orientado al usuario general como por ejemplo las aplicaciones de mensajería.

Por otro lado, en Android *Things* se eliminan las características mencionadas previamente, como por ejemplo algunas APIs y aplicaciones que están normalmente para el uso cotidiano en Android como las aplicaciones de mensajería. Las pantallas opciones de un dispositivo han sido modificadas en sus comportamientos, como por ejemplo notificaciones, interfaz de usuario; en Android *Things* han sido eliminadas. Por otro lado, cuando el desarrollador trabaja utilizando esta plataforma obtienen la biblioteca de *Things* como soporte para el desarrollo. Esto deja en evidencia que la gestión de dispositivos así como la periferia de entrada/salida esta incorporada en Android *Things*. La Figura 2 muestra la arquitectura de Android *Things* que utiliza un sistema de módulo o arquitectura de SOM.

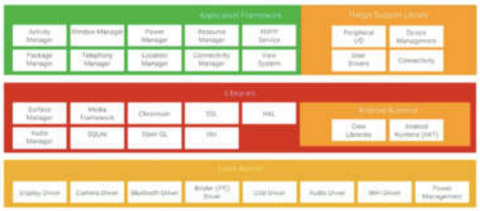


Figura 2: Arquitectura Android Things.

2.2.2 Sistema Operativo

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 [16]. Esto incluye la misma interfaz de usuario, tookit, soporte multimedia y APIs de conectividad utilizadas por los desarrolladores de *Mobile*. Las aplicaciones se integran fácilmente con los servicios populares de Google, como *Firebase* [17], *TensorFlow* [18] y *Google Cloud Plataform* [19] utilizando la diversidad de bibliotecas de Android *Mobile*.

El desarrollo de Android *Things* utiliza el mismo lenguaje y herramientas que se utiliza para desarrollar 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 variantes de servicios de Google Play optimizada específicamente 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.

2.2.3 Actualizaciones

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. Esto permite mantener protegidos a los usuarios en todo momento.

Por otro lado, las imágenes del sistema están firmadas por Google y verificadas para la 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 instalas por el usuario no son soportadas.

Android *Things Console* [20] proporciona herramientas para instalar y actualizar la imagen del sistema en dispositivos de hardware compatibles. Permitiendo enviar actualizaciones a los usuarios así como probar las implementaciones en su propio hardware. La utilización de la consola permite: (i) descargar e instalar la última imagen del sistema Android *Things*; (ii) Crear imágenes de fábrica que contengan aplicaciones OEM junto con la imagen del sistema; (iii) Lanzar las actualizaciones por aire (OTA), incluidas las aplicaciones OEM y la imagen del sistema, a los dispositivos.

2.3 Integración de Android *Things*

Los dispositivos IoT que son desarrollados utilizando Android *Things* cuentan con una placa base de soporte, como por ejemplo el NXP i.MX7D o Raspberry Pi 3 [22]. Para la utilización de estas placas es necesario descargar a la memoria flash o SD Android *Things*. Esta acción permite conectarse a la red Wifi o mediante USB a una computadora. Estas placas posibilitan incorporar hardware adicional conectando a los distintos periféricos accediendo a los protocolos o, a una biblioteca estándar como por ejemplo GPS. Haciendo uso del IDE

Android Studio se inicia la creación del dispositivo IoT.

3. RESULTADOS OBTENIDOS / ESPERADOS

Este trabajo de investigación permitió obtener diferentes resultados que serán utilizados para poder transformar aplicaciones de Android *Mobile* a aplicaciones de Android *Things* confiables e iniciar una metodología o herramienta que automaticen este proceso. Se describen los principales resultados.

Se determinó que Android *Things*: i) permite transformar aplicaciones basadas en Android de una manera transparente y confiable; ii) permite unificar velocidades de procesamiento, minimizar el uso de memoria esto se debe a que Android *Things* ejecuta una única aplicación; iii) en la arquitectura se elimina la capa de aplicación y modifica las APIs, en particular la de la pantalla y la interfaz gráfica; iv) incorpora una consola de actualización gestionada por Google; v) los dispositivos IoT son soportados por una placa base con periféricos de entrada/salida que permiten adaptar hardware adicional conectándose a través de protocolos de comunicación o utilizando la biblioteca estándar.

4. FORMACIÓN DE RECURSOS HUMANOS

Las investigaciones realizadas así como los resultados obtenidos en este trabajo contribuyen al 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 y Ingeniería en Computación de la Universidad Nacional de San Luis, en el marco de los proyectos de investigación.

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), 89-102.

[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 (IJCSEIT)*, *5*(1), 1-5.

[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 low-end devices in the internet of things: a survey. *IEEE Internet of Things Journal*, *3*(5), 720-734.

[12] Wortmann, F., & Flüchter, K. (2015). Internet of things. Business & Information Systems Engineering, 57(3), 221-224.

[13] DiMarzio, J. F. (2015). Setting Up Android Studio. In *Android Studio Game Development* (pp. 1-8). Apress, Berkeley, CA.

[14] Sitio oficial Android Studio: https://developer.android.com/studio/index.html.

[15] PEREZ, Norma Beatriz, et al. Análisis sistematico de la seguridad en internet of things. En XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste). 2018.

[16] 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. 221-233). ACM.

[17] Sitio oficial Firebase: https://firebase.google.com/

[18] Sitio oficial TensorFlow: https://www.tensorflow.org/

[19] Sitio oficial Google Cloud *Plataform*: https://cloud.google.com/

[20] Sitio oficial de desarrollo Android Things: https://developer.android.com/things/console/index.html

[21] Pi, R. (2015). Raspberry Pi Model B.

Aplicaciones de Sistemas de Software: Innovación en procesos, productos y servicios

Hugo Ramón[1], Leonardo Esnaola[2], Juan Pablo Tessore[3], Sebastian Adorno[4], Ignacio Rubio[4]

Instituto de Investigación y Transferencia de Tecnología (ITT)[5]
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento Nro. 1119 3er Piso, Junín (B) – TE: (0236) 4477050 INT 11610

hugo.ramon@itt.unnoba.edu.ar / leonardo.esnaola@itt.unnoba.edu.ar / juanpablo.tessore@itt.unnoba.edu.ar

Resumen

El avance de las tecnologías es acompañado del acrecentamiento de los sistemas de software. En este contexto, la innovación juega un papel muy importante en el desarrollo de productos y servicios. Al incluir al usuario en los procesos de innovación se está dando lugar a la aparición de nuevos espacios de interacción y comunicación en los que el actor que consume productos y servicios deja de considerarse en forma pasiva, para convertirse en un actor activo. Las posibilidades actuales del cómputo ubicuo (o computación psicológicamente invisible), permiten el monitoreo (y cambio) del comportamiento humano y del ambiente donde este se desarrolla. Abarca aspectos que van desde los fundamentos del desarrollo (técnicas de Ingeniería de requerimientos, metodologías de gestión, desarrollo de proyectos, técnicas de planificación, métricas, normas de calidad, redes de sensores y algoritmos inteligentes y procesamiento de señales) hasta la concepción de aplicaciones específicas que impactan en los procesos

---

1 Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC

2 Docente Investigador ITT / Doctorando UNLP

3 Docente Investigador ITT / Becario Doctoral CIC

4 Becario ITT

5 ITT - Centro Asociado CIC

productivos, los procesos de gobierno, las ciudades digitales y la educación.

Palabras clave:

Innovación, Inteligencia artificial, *e-services*.

Contexto

Las líneas de investigación presentadas en este trabajo se enmarcan en el proyecto "Tecnología y aplicaciones de Sistemas de Software: Innovación en procesos, productos y servicios" presentado en el marco de la convocatoria a Subsidios de Investigación Bianuales (SIB2019) de la Secretaría de Investigación, Desarrollo y Transferencia de la UNNOBA. 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 y Transferencia en Tecnología (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.

Introducción

Se define una tendencia tecnológica estratégica como aquella con potencial innovador sustancial [1], que empieza a salir de un estado de emergente hacia un impacto de más amplio uso; o aquellas tecnologías que tienen rápido crecimiento pero alta volatilidad y pueden alcanzar su punto de inflexión en los próximos años.

La empresa experta en consultoría e investigación de las tecnologías de la información, Gartner Inc., enunció que las tecnologías estrategias para 2019 se pueden resumir en:

- Cosas Autónomas: Interacción más natural con el ambiente.
- Analitica Aumentada: Utilización de algoritmos automatizados para explorar nuevas hipótesis desde los datos.
- AI Driven Development: estudio de herramientas, tecnologías y mejores prácticas para integrar la IA en las aplicaciones y en las mismas herramientas para el proceso de desarrollo.
- Gemelos Digitales. Un gemelo digital es un modelo de software dinámico de un objeto físico o un sensor de datos que entiende el estado, responde a los cambios, mejora las operaciones y agrega valor.
- Edge Computing. El procesamiento de la información y la recopilación y entrega de contenidos se sitúan más cerca de las fuentes de la información.

- Tecnologías de Inmersión. cambiar la forma en que los usuarios interactúan con el mundo, tecnologías como la realidad aumentada (RA), la realidad mixta (RM) y la realidad virtual (RV) aplican a esto.
- Blockchain. Blockchain es un tipo de libro distribuido, una lista ordenada cronológicamente y en expansión de registros de transacciones firmados criptográficamente, irrevocables y compartidos por todos los participantes en una red.
- Smart Spaces. Un espacio inteligente es un entorno físico o digital, en el que los seres humanos y los sistemas tecnológicos interactúan en ecosistemas, cada vez más abiertos, conectados, coordinados e inteligentes.
- Ética Digital y Privacidad. Los consumidores son cada vez más conscientes del valor de su información personal y están cada vez más preocupados por la forma en que está siendo utilizada por entidades públicas y privadas.
- Quantum Computing. La computación cuántica es un tipo de computación no clásica que se basa en el estado cuántico de las partículas subatómicas que representan información como elementos denotados como bits cuánticos o "qubits".

Estas tecnologías estratégicas impactan en cómo se desarrollan aplicaciones, buscando lograr la integración que las mismas deben implementar para descubrir las necesidades de los usuarios, con el fin de presentar la información pertinente en el lugar correcto y en el momento adecuado, con las características de calidad requeridas.

Estas características intrínsecas de los nuevos sistemas hacen necesario el uso de diferentes técnicas de ingeniería de software, entre otras, el desarrollo orientado a servicios, desarrollo basado en componentes, plugins basados en arquitecturas, sistemas basados en eventos, microservicios, y la evolución del software dinámico, todo basado en herramientas inteligentes.

La oportunidad de I+D+i se da en adaptar/diseñar aplicaciones o a la forma de construir las mismas, las cuales gestionan información procedente de redes de sensores compuestas por diversos nodos distribuidos y localizadas en espacios cerrados (hospitales, fábricas, oficinas, etc.) o abiertos (campos, bosque, etc.). Ambos espacios se encuentran en las diferentes actividades económicas de la región de influencia de esta universidad.

Líneas de Investigación, Desarrollo e Innovación

Se han identificado dos líneas de investigación principales: “Gestión de la Innovación”: procesos metodológicos para abordar Innovación y su implicación medible en procesos productivos y “*Smart Places* en Gobiernos/Empresas/ Ciudades/Universidades”: Investigar / Seleccionar / Adaptar / Desarrollar

herramientas inteligentes para el desarrollo de aplicaciones y/o la simulación de ambientes inteligentes para mejoras de los procesos productivos.

Resultados y Objetivos

Uno de los objetivos principales de este proyecto es determinar cómo se modifican los procesos de desarrollo para smart places con la incorporación del usuario final (como partícipe o representante de la sociedad). Dicho objetivo será abordable mediante dos objetivos específicos:

- Gestión de la Innovación
- Smart places Gobierno / Empresa / Ciudades / Universidades

A su vez, se busca promover la formación de grado/postgrado (Maestrías y Doctorados) de los miembros de proyecto. También fomentar la Interacción con otros grupos de I+D+i del país y del exterior en la temática del proyecto. Finalmente difundir y transferir los logros alcanzados mediante la presentación y participación en diferentes revistas especializadas, congresos, jornadas y workshops, todos nacional es/internacionales relacionados con la temática de estudio.

Formación de Recursos Humanos

Se espera que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su rol de investigadores o partícipes activos en equipos de investigación, transferencia e innovación; fomentando la culminación de sus estudios superiores, promoviendo la redacción, exposición y defensa de Trabajos Finales de Grado y Postgrado, y la realización de Prácticas Profesionales Supervisadas.

En relación a este tema, para los próximos dos años se espera contribuir al inicio y concreción de 2 (dos) Tesinas de Licenciatura en Sistemas, 3 (tres) Prácticas Profesionales Supervisadas de la Ingeniería en Informática, 1 (una) Tesis de Especialización, 1 (cuatro) Tesis de Magíster y 2 (dos) Tesis Doctorales.

Bibliografía

Adams, C. G. (2018). Agent And Component Object Framework For Concept Design Modeling Of Mobile Cyber Physical Systems. NAVAL POSTGRADUATE SCHOOL MONTEREY CA MONTEREY United States.

Atkinson, R. (1999). Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria. International journal of project management, 17(6), 337-342.

Carroll, J. M. (2019). The Internet of Places. In *Social Internet of Things* (pp. 23-32). Springer, Cham.

Ciarletta, L., Leclerc, T., Siebert, J., Chevrier, V., & Schaff, A. (2008). Towards standards for Pervasive Computing evaluation: using the

multi-model and multi-agent paradigms for mobility.

Deane, R. H., Clark, T. B., & Young, A. P. (1997). Creating a learning project environment: aligning project outcomes with customer needs. Information systems management, 14(3), 54-60.

Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. Human-computer interaction, 16(2), 97-166.

Fujinami, K., Yamabe, T., & Nakajima, T. (2004, November). Bazaar: a conceptual framework for physical space applications. In International Symposium on Ubiquitous Computing Systems (pp. 174-191). Springer Berlin Heidelberg.

Glover, I., & McDonald, K. (2018, June). Digital places: location-based digital practices in higher education using Bluetooth Beacons. In *EdMedia+ Innovate Learning* (pp. 950-959). Association for the Advancement of Computing in Education (AACE).

Gračanin, D., Eltoweissy, M., Cheng, L., & Tasooji, R. (2018, July). Reconfigurable spaces and places in smart built environments: A service centric approach. In *International Conference on Human-Computer Interaction* (pp. 463-468). Springer, Cham.systems: Current challenges and future networking applications. IEEE Access, 6, 12360-123

Guo, Y., Hu, X., Hu, B., Cheng, J., Zhou, M., & Kwok, R. Y. (2018). Mobile cyber physical

Harter, A., & Hopper, A. (1994). A distributed location system for the active office. IEEE network, 8(1), 62-70.

Ika, L. A. (2009). Project success as a topic in project management journals. Project Management Journal, 40(4), 6-19.

Kuniavsky, M. (2010). Smart things: ubiquitous computing user experience design. Elsevier.

Lee, Jay; Bagheri, Behrad; Kao, Hung-An (2014). "Recent Advances and Trends of Cyber-Physical Systems and Big Data Analytics in Industrial Informatics". IEEE Int. Conference on Industrial Informatics (INDIN) 2014.

López, O., Blanco, M., & Guerra, S. (2017). Evolución de los modelos de la gestión de innovación. *Innovaciones de negocios*, (10).

Modahl, M., Agarwalla, B., Abowd, G., Ramachandran, U., & Saponas, T. S. (2004, October). Toward a standard ubiquitous computing framework. In Proceedings of the 2nd workshop on Middleware for pervasive and ad-hoc computing (pp. 135-139). ACM.

Motta, J. J., Zavaleta, L., Llinás, I., & Luque, L. (2013). Procesos de innovación y competencias de los recursos humanos

en la industria del software en Argentina. Revista Iberoamericana de Ciencia, Tecnología y Sociedad, 24(8), 145-145.

Serra, A. (2014). Tres problemas sobre los laboratorios ciudadanos: Una mirada desde Europa. Revista iberoamericana de ciencia tecnología y sociedad, 8(23), 283-298.

Weiser, M. (1993). Hot topics-ubiquitous computing. Computer, 26(10), 71-72.

Aplicaciones Web Progresivas Impulsadas por el Avance de los Estándares Web

Rocío A. Rodríguez[1,2], Pablo M. Vera[1,2], M. Roxana Martínez[1], Fernando A. Parra Beltrán[1], Artemisa Trigueros[2], Mariano G. Dogliotti[2]

[1]CAETI - Centro de Altos Estudios en Tecnología Informática
Universidad Abierta Interamericana (UAI)
Montes de Oca 745, Ciudad Autónoma de Buenos Aires, Argentina

[2]GIDFIS – Grupo de Investigación, Desarrollo y Formación en Innovación de Software
Departamento de Ingeniería e Investigaciones Tecnológicas
Universidad Nacional de La Matanza

{rocioandrea.rodriguez, pablomartin.vera, roxana.martinez}@uai.edu.ar
{Adolfo.ParraBeltran} @alumnos.uai.edu.ar
{atrigueros, mdogliotti}@unlam.edu.ar

RESUMEN

La tecnología móvil ocupa un papel primordial en la vida de las personas hasta el punto por el cual en la actualidad el teléfono móvil combinado con internet es el principal medio por el cual se accede a la información. La mejora permanente en el hardware de los Smartphone permite disponer de equipos cada vez más poderosos que incluyen una gran cantidad de sensores (como ser GPS, de luz, ritmo cardiaco, de temperatura, presión atmosférica, etc). Estos sensores permiten a las personas interactuar con el medio. El propósito de esta línea de investigación, es el análisis de los nuevos estándares que permiten a las aplicaciones web hacer uso del hardware, esto impulsa a la construcción de aplicaciones web, ganando en portabilidad en comparación con el desarrollo nativo. Pero es importante que el usuario no sienta que ha perdido ningún elemento o atractivo que podía tener en una aplicación nativa, por ejemplo contar con un ícono para abrir la aplicación, recibir notificaciones, etc. Estas características se pueden incorporar en las aplicaciones web utilizando el enfoque PWA (Aplicaciones Web Progresivas), uno de los principales puntos de estudio de este artículo.

Palabras clave: Web Móvil, Aplicaciones Web, Aplicaciones Nativas, Dispositivos Móviles, API, W3C, PWA

CONTEXTO

Esta línea de I+D forma parte de los proyectos radicados en el Centro de Altos Estudios en Tecnología Informática (CAETI) de la Universidad Abierta Interamericana (UAI). En este proyecto participan docentes y alumnos tanto de sede Centro como de la Castelar (ambas en la provincia de Buenos Aires). El proyecto cuenta con financiamiento asignado y una duración de 2 años. Analizándose los estándares del W3C que facilitan el acceso del hardware desde los navegadores. La participación activa de los alumnos contribuye a poder trabajar en paralelo sobre diversos estándares.

Este artículo se realiza en conjunto con la Universidad Nacional de La Matanza en donde el grupo de investigación GIDFIS tiene una línea de investigación y desarrollo también en el área de dispositivos móviles vinculado con el enfoque PWA (Aplicaciones Web Progresivas). Siendo este un proyecto PROINCE (Programa de Incentivos) que inició este año continuando una línea previa de I+D y tiene una duración de 2 años.

La colaboración entre ambos equipos se facilita por 2 docentes investigadores que forman parte de ambas instituciones y esto nutre a ambos equipos, dando un enfoque integral en cuanto a software y hardware.

1. INTRODUCCIÓN

La web móvil cada vez se hace más presente para poder buscar información, es que entre otros dispositivos móviles es el teléfono celular el medio de consulta más cercano. No obstante, los usuarios móviles tienen diversas aplicaciones instaladas en sus dispositivos, las que permiten en algún caso acceder al hardware (por ejemplo, para posicionarse haciendo uso del GPS), es allí donde cobran vital importancia los estándares del W3C (consorcio web a nivel internacional) [1]. Estos estándares posibilitan que los distintos browsers permitan el acceso al hardware de los dispositivos desde la web. "La brecha entre las capacidades de las aplicaciones nativas sobre las aplicaciones web se está reduciendo, especialmente impulsada por los estándares web de acceso al hardware". [2].

Entonces, el acceso limitado al hardware que antes se tenía en una aplicación web, deja de ser un problema. Esto sumado a la portabilidad de las aplicaciones web y la posibilidad de trabajar offline, hace que ya no exista una brecha que distancie las aplicaciones web de las aplicaciones nativas.

La constante evolución de los dispositivos móviles con hardware cada vez más especializado merece poner el foco en ~~de~~ facilitar por medio de APIs el manejo de sensores y otros componentes (ver figura 1). A demás de incorporar cada vez más componentes estos van mejorando en prestaciones, si se considera por ejemplo la cámara fotográfica, actualmente la mayoría de los modelos dispone de 2 cámaras, una trasera y otra frontal, pero además hay modelos que incorporan para tomar una foto más de una cámara, no todas funcionan igual por ejemplo se puede contar con: (a) de dos lentes iguales, lentes con distintas distancias focales, (b) una lente común y otra gran angular, (c) una lente color y otra blanco y negro, (d) un sensor común y otro de profundidad. Los avances tecnológicos van impactando en todos los componentes, actualmente se cuenta con smartphones por ejemplo de 8 núcleos de distintas velocidades, los cuales se aprovechan para distintas tareas. No sólo se debe considerar la CPU sino que tienen también GPU, lo cual aumenta significativamente la potencia y posibilidad de correr mayor tipo de aplicaciones con una muy buena performance.

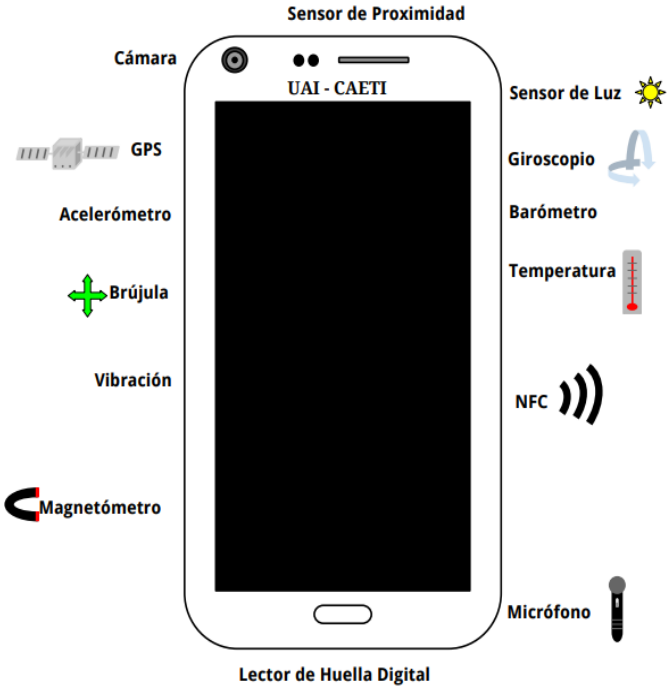


Figura 1. Componentes de Hardware – Smartphones

Cabe destacar que no todos los smartphones traen estos componentes o sensores, por lo que existen diversas aplicaciones que permiten testear los componentes disponibles en un dispositivo. De hecho, en muchos casos un dispositivo de una determinada marca el mismo modelo dependiendo del país en el que se compre tiene equipamiento distinto (por ejemplo se puede mencionar el celular Motorola G7 Plus en Argentina se vende sin NFC [3], sin embargo en España tiene NFC integrado).

En la figura 2 se presenta la captura de pantalla en donde pueden visualizarse los sensores presentes en un dispositivo en este caso un Motorola G6. Como puede observarse, la aplicación de testeo muestra un listado de componentes aquellos con nombre

grisados son los que no incorpora el dispositivo. Cada aplicación de testeo a su vez incluye su propia lista de componentes, por ejemplo, en la figura 2 no se muestra NFC que en particular este dispositivo de todos modos no cuenta con el chip NFC. Con lo cual si se quiere verificar un sensor en particular deberá asegurarse que esté en el listado de la aplicación de testeo.

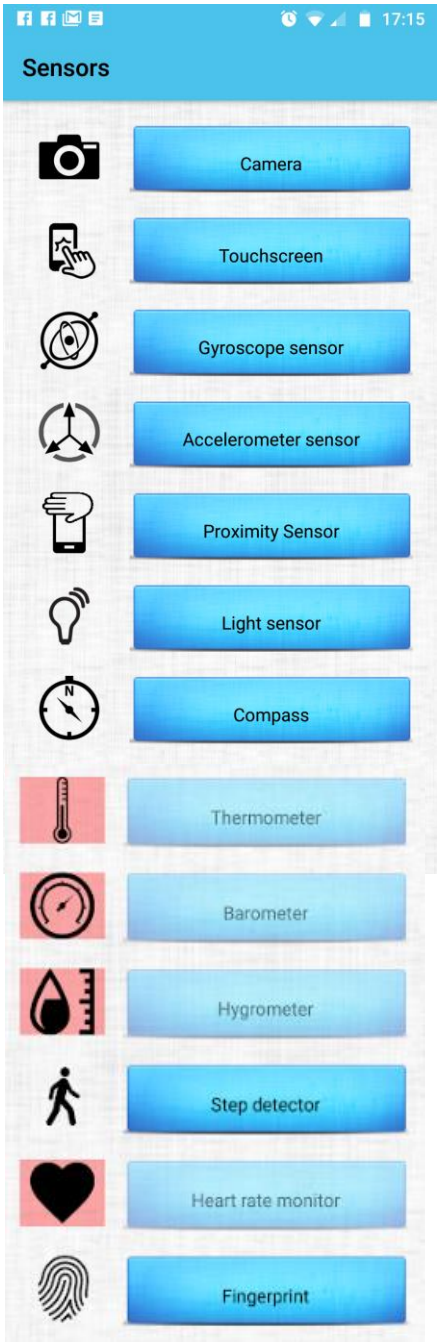


Figura 2. Aplicación de Testeo de Sensores

En cuanto a software se evidencian aún diferencias, a pesar de que existen frameworks que facilitan el diseño de las aplicaciones web móviles dándoles un aspecto verdaderamente idéntico a los controles nativos de un determinado sistema operativo. Una de las diferencias es poder tener un ícono que permita acceder a la aplicación. Es por lo que aparece el enfoque PWA.

Para acercar más a las aplicaciones web de las nativas "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" [4].

"Las aplicaciones web progresivas son una evolución natural de las aplicaciones web que difuminan la barrera entre la web y las aplicaciones, pudiendo realizar tareas que generalmente solo las aplicaciones nativas podían llevar a cabo. Algunos ejemplos son las notificaciones, el funcionamiento sin conexión a Internet o la posibilidad de probar una versión más ligera antes de bajarte una aplicación nativa de verdad" [5].

Resulta importante tener un enfoque integral que conjugue aspectos de hardware y software. "Una de las características importantes de la gran mayoría de los desarrollos móviles es su corta duración. Esto se debe a factores como la gran competencia en el sector, los cambios en el mismo con la aparición, casi constante, de novedades tanto software como hardware" [6].

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.
- Facilitar el uso de las aplicaciones web móviles, reduciendo las diferencias que presentan en cuanto a software con respecto a las aplicaciones nativas.

3. RESULTADOS OBTENIDOS/ESPERADOS

En el CAETI se han construido aplicaciones nativas (ANDROID) y web, equivalentes, que permitieron poder comprobar que la performance de la web es incluso superior en algunas acciones particulares a las aplicaciones nativas [2]. Es decir que la web no pierde performance comparada con una aplicación instalable. Por otra parte, la incorporación de hardware puede hacerse presente en la mayoría de las aplicaciones dándoles valor agregado, por ejemplo, utilizar el sensor de proximidad para realizar gestos que desencadenen acciones [7] ó bien que el teléfono vibre podría dar cierto feedback ante un evento concreto, lo cual se puede lograr haciendo uso de la API de vibración [8]. En esta línea de investigación se analizan estándares más bien vinculados con el hardware del dispositivo móvil para poder utilizados desde la web. Dando por resultado comparativas de tiempos de respuesta de sensores en diversos dispositivos, así como también APIs que solucionan el manejo de dichos sensores en aplicaciones futuras. Estas futuras aplicaciones serán abordadas bajo el enfoque PWA permitiéndole al usuario tener una aplicación con las mismas prestaciones que una nativa (en cuanto a interfaz, ícono de acceso, etc).

4. FORMACIÓN DE RECURSOS HUMANOS

En UAI el grupo de investigación está formado por 8 personas:

- 3 Docentes (2 de Postgrado y 1 de Grado).
- 5 (Alumnos de Posgrado y Grado).

En UNLaM el grupo de investigación está formado por 11 personas.

- 10 Docentes (Grado).
- 1 (Alumnos Grado).

En el área de dispositivos móviles se encuentran en realización 3 tesis de maestría (1 en la UAI y 2 en UNLaM). También se encuentra en realización 1 tesis de doctorado en la UNLP (Universidad Nacional de La Plata).

5. BIBLIOGRAFIA

[1] "W3C. "About W3C". 2019 https://www.w3.org/Consortium/

[2] Rodríguez Rocío A, Vera Pablo M, Martínez M. Roxana, Pons Claudia, Valles Federico E, de La Cruz Luis Verbel. "Reducing the Gap between Native and Web Applications". SETESEC – Italia2014

[3] NFC FORUM, About the Technology NFC and Contactless Technologies https://nfc-forum.org/what-is-nfc/about-the-technology/

[4] Thomas Pablo, Delia Lisandro, Corbalán Leonardo, et al. “Tendencias en el desarrollo de Aplicaciones para Dispositivos Móviles”. Instituto de Investigación en Informática LIDI (III-LIDI) Facultad de Informática – Universidad Nacional de La Plata. Centro Asociado a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC). Abril 2018. https://digital.cic.gba.gob.ar/bitstream/handle/11746/8316/11746_8316.pdf?sequence=1&isAllowed=y

[5] Ramírez Ivan, “¿Qué es una aplicación web progresiva?”. Enero 2019 https://www.xataka.com/basics/que-es-una-aplicacion-web-progresiva-o-pwa

[6] Vique, R. R. (2012). Métodos para el desarrollo de aplicaciones móviles.

[7] Vera, Pablo Martin, Rocio Andrea Rodriguez. "Creating and Using Proximity Events on Mobile Websites." IEEE Latin America Transactions 14.11 (2016): 4579-4584.

[8] W3C, Vibration API (Second Edition), Octubre 2016 https://www.w3.org/TR/2016/REC-vibration-20161018/

AVANCES EN EL PROYECTO ROBÓTICA BASADA EN INTERNET DE LAS COSAS

Azcurra, D., Santos, D., Fernández, G., Fernández, S., Higa, E., Amaro, M., Wieilly, A.

Dpto. Desarrollo Productivo y Tecnológico
Universidad Nacional de Lanús.
29 de Septiembre 3901 (1826) Remedios de Escalada, Lanús
Buenos Aires, Argentina.

Resumen

Este proyecto continúa 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 actividades, logros y objetivos alcanzados a la fecha con relación 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.

Luego del primer año de ejecución del proyecto, y según lo planificado, se completaron las etapas:

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.

y se comenzó con la:

III. Identificar las modificaciones necesarias para adaptar los mecanismos de administración y control de robots para que interactúen con plataformas de IoT.

Para el próximo año está previsto continuar con las actividades anteriores y:

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 y tres alumnos avanzados de la carrera de Licenciatura en Sistemas de la Universidad Nacional de Lanús.

Atendiendo a lo previsto en cuanto a transferencia a la actividad docente, vale la pena mencionar las siguientes actividades desarrolladas durante el primer año de proyecto:

•Se capacitó a los alumnos integrantes del equipo en la metodología de trabajo para que puedan experimentar y desarrollar capacidades en las áreas de arquitectura de computadoras, sistemas embebidos, robótica, automatización, comunicaciones, sistemas distribuidos, programación concurrente y procesamiento de datos en tiempo real.

•En el marco del proyecto, uno de los alumnos desarrolló su Trabajo Final Integrador (está esperando la fecha para defenderlo) y los otros dos los están desarrollando.

•Se realizaron dos presentaciones en el marco de la Semana del Estudiante los días 27 y

28 de Marzo de 2018 en representación de la carrera de Licenciatura en Sistemas.

•Se colaboró con el stand de la carrera de Licenciatura en Sistemas en la Expo Carreras 2018.

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/

[6] Azcurra, D., Santos, D., Fernández, G., Fernández, S., (2018), *"ROBÓTICA BASADA EN INTERNET DE LAS COSAS"*. Proceedings del XX Workshop de Investigadores en Ciencias de la Computación. Pág. 835 – 837. ISBN 978-987-3619-27-4.

[7] Kranz, M. (2017). *"Building the Internet of Things"*. ISBN 978-1-119-28566-3. Ed. Wiley.

[8] Sarangi, S., R., Sethi, P. (2016). *"Internet of Things: Architectures, Protocols, and Applications"*

[9] Al-Fuqaha, Ala., Guizani, Mohsen., Mohammadi, Mehdi., Aledhari, Mohammed., Ayyash, Moussa. (2015). *"Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications"*.

[10] McEwen, Adrian., Cassimally, Hakim. (2014), *"Designing the Internet of Things"*. ISBN 978-1-118-43063-7. Ed. Wiley.

Azcurra, D., Rojo, S., Rodríguez, D., (2013), "AVANCES EN EL PROYECTO ARQUITECTURAS DE CONTROL PARA ROBOTS AUTÓNOMOS MÓVILES DIDÁCTICOS BASADAS EN SISTEMAS EMBEBIDOS", Proceedings del XV Workshop de Investigadores en Ciencias de la Computación. Pág. 835 – 837. ISBN 978-987-3619-27-4. Pág. 987-989. ISBN: 9789872817961

Sitios web:

IBM. *"IaaS PaaS SaaS – Modelos de servicio Cloud": https://www.ibm.com/cloud-computing/es-es/learn-more/iaas-paas-saas/*

http://tinyos.stanford.edu/tinyos-wiki/index.php/FAQ#What_is_TinyOS.3F

https://riot-os.org/#features

http://www.contiki-os.org/#why

https://thinger.io/

https://www.freertos.org/about-RTOS.html

https://thingsboard.io/docs/getting-started-guides/what-is-thingsboard/
https://devicehive.com/

MyDevices Cayenne: https://mydevices.com/
AWS IoT Core "What is AWS IoT?": *https://docs.aws.amazon.com/iot/latest/developerguide/what-is-aws-iot.html*

Microsoft Azure. (17/05/2018) "Introduction to Azure and the Internet of Things". Microsoft Docs: *https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-introduction*

Microsoft Azure. (17/05/2018) "Azure IoT technologies and solutions: PaaS and SaaS". Microsoft Docs: *https://docs.microsoft.com/en-us/azure/iot-fundamentals/iot-services-and-technologies*

Google Cloud "Google Cloud IoT": *https://cloud.google.com/solutions/iot/?hl=es*

IBM Cloud. (08/05/2018). "Watson IoT Platform Feature Overview": *https://console.bluemix.net/docs/services/IoT/feature_overview.html#feature_overview*

Computación en la nube para un sistema de monitoreo de calidad de aire y ruido ambiental

Zaradnik, Ignacio; Lupi, Daniel; Agüero, Agustin; Behar, Christian; Lanzilliotti, Leandro; Vázquez, Matías; Canziani, Mónica.
Laboratorio de Inteligencia Ambiental Departamento de Ingeniería e Investigación Tecnológica, Universidad Nacional de la Matanza. Buenos Aires, Argentina
agustinaguero.aa@gmail.com; behar.christian@yahoo.com.ar; lanzillottilea@yahoo.com.ar; mati.pity@gmail.com; izaradnik@unlam.edu.ar

RESUMEN

El presente trabajo detalla el uso de Computación en la Nube (Cloud Computing) para un sistema de monitoreo de calidad de aire y ruido ambiental. Se comienza con una introducción a la temática de ciudades inteligentes y su importancia para el futuro. A continuación, se presenta el diagrama general del sistema y breve una explicación del funcionamiento de este, en donde se detallan los principales elementos. Y se finaliza con descripción del uso del Portal de uno de los tantos proveedores de computación en la nube.

Palabras Clave: Ciudades inteligentes, Calidad de aire, Ruido ambiental, Computación en la Nube.

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][4]. El presente trabajo es parte de lo realizado en el marco del proyecto "Internet de las cosas y sus aplicaciones en ciudades inteligentes", el cual se viene desarrollando desde el 2018. Este trabajo se desarrolla con fondos provenientes del programa de Investigaciones PROINCE.

1. INTRODUCCION

Las ciudades modernas son las responsables del 80% de la producción económica mundial y del 70% del consumo de energía mundial y las emisiones de gases de efecto invernadero, y, para el año 2050 serán el hogar del 66% de los 9.000 millones de habitantes proyectados. Por lo tanto, es necesario pensar en políticas que permitan la sustentabilidad de las ciudades [5]. Las ciudades sustentables deben mejorar la calidad de vida de los habitantes, la eficiencia de las operaciones y los servicios urbanos, y su competitividad, mientras se asegura que esta se encuentre de acuerdo a las necesidades de las generaciones presentes y futuras en lo que respecta a aspectos económicos, sociales, medioambientales así como también culturales. Las tecnologías de la información y las comunicaciones, y en especial Internet de las cosas, tienen un papel muy importante, estas actúan como una plataforma para obtener datos e información que ayudará a mejorar en entendimiento sobre como la cuidad está funcionando en termino de consumo de recursos, servicios y calidad de vida. Como ejemplo de los problemas cotidianos que se intentan solucionar con estos dispositivos podemos citar la potencial aparición mosquitos a raíz de un aumento en la temperatura y humedad en un parque, la cual puede ser detectada por sensores y generar un proceso de fumigación [6]. Otro ejemplo es la red de monitoreo de calidad de aire y ruido de la ciudad de Buenos Aires, la cual permite controlar la calidad ambiental del entorno urbano a través de un seguimiento continuo de los niveles de contaminación y así generar información confiable, comparable y representativa para su aplicación en la estrategia local de protección de la salud y el ambiente [7].

2. LINEAS DE INVESTIGACION y DESARROLLO

El objetivo general de este trabajo es desarrollar, implementar y estudiar los

resultados de uso de sistemas experimentales asociados a la temática de ciudades inteligentes, como ser estacionamiento inteligente, control de polución, control de tráfico, entre otras. Para ello se investigará: las distintas aplicaciones, sus influencias en la comunidad y la posibilidad de implementar en la Universidad; la arquitectura de las aplicaciones, tanto a nivel hardware como software; las distintas tecnologías de sensado, control y comunicación empleadas.

3. RESULTADOS OBTENIDOS

3.1. Descripción General

La figura N°1 presenta el diagrama en bloques del sistema de monitoreo de calidad de aire y ruido ambiental. El corazón del sistema es un microcontrolador de 32 bits, STM32F411, de la firma ST. Este es el encargado de efectuar la lectura de los sensores, procesar y analizar los datos obtenidos y transmitirlos, a través del módulo inalámbrico, a un servidor en la nube.

Figura N°1. Diagrama en bloques del sistema.

El procesado implica corregir las variaciones de los datos medidos debido a la temperatura, mientras que el análisis verifica, en forma local, si los datos excedan los valores de limites preestablecidos. Si los datos exceden estos limites, se transmitirá de forma inmediata un aviso al Portal de datos y se activará una indicación en forma local. Caso contrario los datos serán transmitidos periódicamente. El módulo inalámbrico empleado es uno de tecnología 3G, UL865, de la empresa Telit. Los sensores utilizados son: temperatura, humedad, ruido ambiental, dióxido de carbono (CO2) y material particulado de hasta 10um. El módulo GPS no se encuentra implementado en la versión actual. La figura N°2 muestra una imagen del sistema de monitoreo de calidad de aire y ruido ambiental experimental desarrollada.

Figura N°2. Sistema experimental.

Para el desarrollo se empleó como proveedor de Computación en la Nube a la empresa Telit, la cual ofrece una Plataforma como un Servicio (PaaS - Platform as a Service) para aplicaciones de Internet de las Cosas (IoT) llamada Telit IoT Portal [8]. Si bien este servicio es pago, el proveedor ofrece la posibilidad de generar una cuenta del tipo demo, opción que se utilizó para este proyecto. El Telit IoT Portal posee entre los mecanismos de transporte los protocolos MQTT, COaP y HTTP [9], pero gracias al uso del módulo UL865 no fue necesario la implementación de ninguno de ellos, ya que el módulo permite el envío de datos a través de simples comandos AT

3.2. Desarrollo

El primer paso realizado fue la generación de una cuenta, a través del link: https://portal-dev.telit.com/app/signup. Luego que esta fuera aprobada, se ingresó al Portal por medio del link: https://portal-dev.telit.com. Dentro del mismo se obtuvieron los datos para la conexión de nuestro módulo al Portal: la dirección a la

cual se debe conectar el módulo (API Server), a través del menú “Desarrolladores” y el “Default Application Token” de la aplicación por defecto, la cual usaremos como base de nuestro proyecto, por medio de la opción “Aplicaciones” del mismo menú. En las figuras N°3 y N°4 se pueden ver los datos seleccionados.

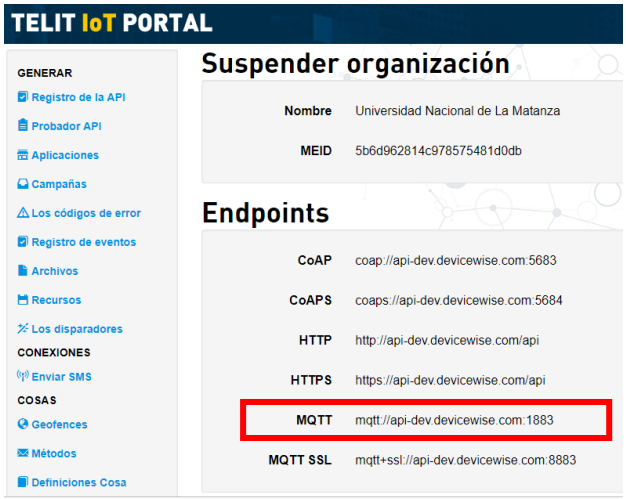


Figura N°3. Dirección de conexión.

Figura N°4. Token de la aplicación.

Luego que se obtuvieron estos datos, se realizó la conexión entre el módulo y el Portal a través de los siguientes comando AT [10]:

- AT+CGDCONT, define el contexto de la conexión, es decir el protocolo de red (IPV4, IPV6), y el APN, el nombre del punto de acceso, de la operada de telefonía celular empleada.
- AT#SGACT, activa el contexto. Con este comando se establece una comunicación con el Gateway de la operadora.
- AT#DWCFG, configuración de la conexión al Portal. En este comando se especifican la dirección del Portal y un código de seguridad, Token.
- AT#DWCONN, conexión al Portal.
- AT#DWSTATUS, chequeo del estado de la conexión al Portal.

El Portal permite que un usuario tenga varias aplicaciones, por ejemplo: nuestro sistema de monitoreo de calidad de aire y ruido ambiental, un sistema de estacionamiento inteligente, un sistema de control de tráfico, etc. Cada aplicación puede tener asociada un conjunto de “Cosas”, pudiendo ser, para los ejemplos anteriores, estaciones de monitoreo de calidad de aire y ruido ambiental, sensores de estacionamiento y semáforos respectivamente. A su vez cada “Cosa” tiene atributos que son las propiedades y las alarmas. En nuestro caso, los valores de los distintos sensores utilizados se relacionaron con las propiedades, mientras que los valores limites prestablecidos con las alarmas. Las propiedades pueden ser creadas manualmente, a través del Portal, o automáticamente, cuando se recibe una solicitud de publicación. En nuestro caso las propiedades fueron creadas automáticamente. En cambio, las alarmas deben crearse manualmente, acción que se realizó.

Para la publicación de los valores de los sensores y el estado de las alarmas de utilizó en comando AT#DWSEND. En el mismo, el primer parámetro representa el tipo de mensaje que se enviara, en nuestro caso un tipo normal. El segundo parámetro representa que acción se realizará, siendo en para nuestra aplicación una publicación de una propiedad o una publicación de alarma. Los siguientes parámetros se toman de a pares siendo, en el caso de los sensores, nombre del sensor y valor y, en el caso de la alarma, nombre de la alarma y estado.

El Portal crea un gráfico por cada una de las propiedades y lo actualiza a mediada que los datos son publicados, en nuestro caso fueron creados 5 gráficos, uno por cada sensor. La figura N°5 presenta la pantalla principal al ingresar al portal, donde se puede ver todas las

“Cosas” asociadas a dicha aplicación y un registro de los eventos. La figura N°6 muestra la pantalla asociada a un “Cosa” particular, previamente seleccionada en la pantalla principal. En esta se puede ver e estado de las alarmas asociadas y gráficos con las distintas propiedades.

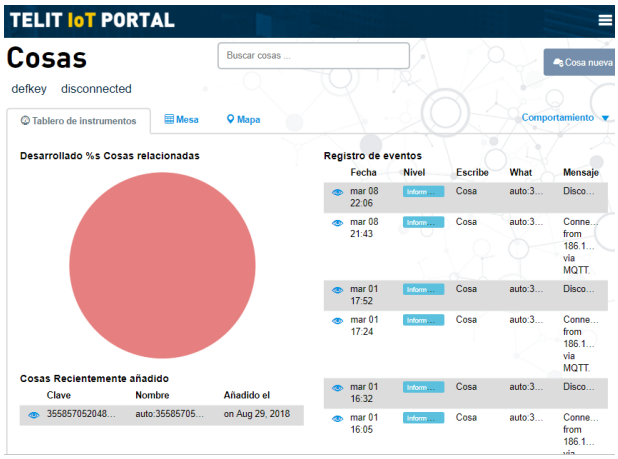


Figura N°5. Pantalla principal del Portal.

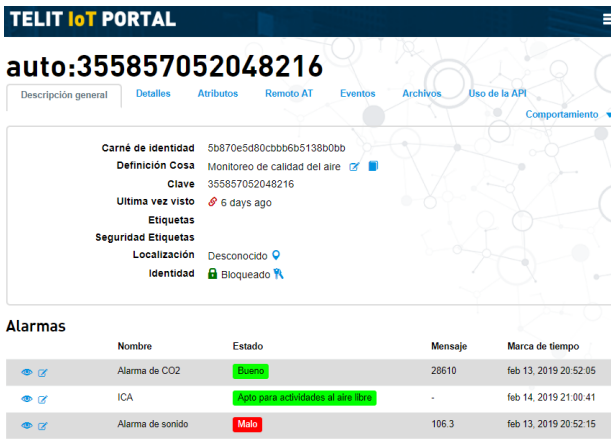


Figura N°6. Pantalla del sistema de monitoreo.

Las figuras N°7 y N°8 presentan los gráficos de los sensores de CO2 y de nivel de ruido.

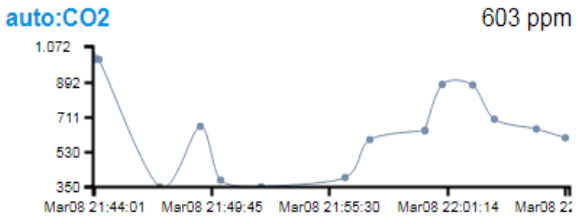


Figura N°7. Representación gráfica de los valores del sensor de CO2.

El Portal permite ver el histórico de los datos de cualquiera de sus propiedades (sensores en nuestro caso), la figura N°9 presenta el histórico del sensor de temperatura.

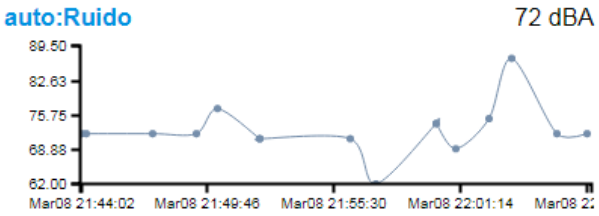


Figura N°8. Representación gráfica de los valores del sensor de nivel de ruido.

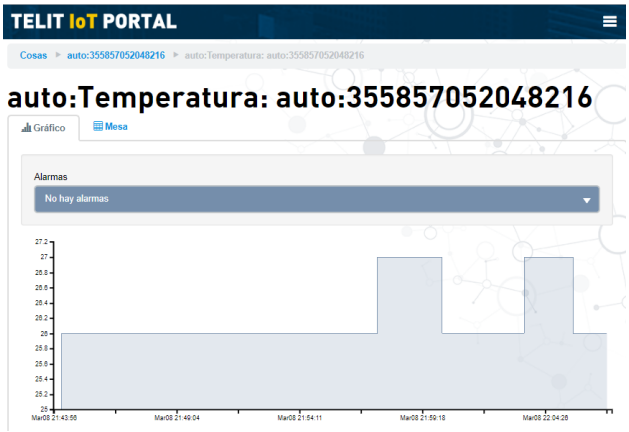


Figura N°5. Representación histórica de los valores del sensor de temperatura.

3.3. Conclusiones

Se logró el desarrollo de un sistema experimental asociado a la temática de ciudades inteligentes, un sistema de monitoreo de calidad de aire y ruido ambiente. En lo que respecta concretamente al tópico de este trabajo, el uso de Computación en la Nube, se logró: comprender distintos aspectos asociados a la Computación en la Nube (tipos de servicios, protocolos relacionados, etc), el uso efectivo del Portal y la generación de una interfaz de visualización de datos y alarmas. Como siguientes pasos están previsto la implementación del sistema desarrollado en el ámbito de la Universidad Nacional de la Matanza, el estudio de datos relevados y la información generada con ellos y el desarrollo de una consola (dashboard) para una mejor visualización de los datos de la estación de monitoreo de calidad de aire y ruido ambiental.

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. La formación grupal busco generar conocimiento en la tematica de ciudades inteligentes y como sus aplicación puede mejorar la vida de los ciudanos. Los casos de la formacion individual se enumeran a continuación: Ignacio Zaradnik la gestión de grupos de trabajos, Monica Canziani la revison bibligrafica y la elaboración de estados del arte, Agustin Agüero el diseño de Consol de datos en la nube, Matias Vazquez el diseño de aplicaciones de sistemas embebidos y tanto Christian Behar como Leandro Lanzilliotti el desarrollo del hardware.

5. REFERENCIAS

[1] Brengi; Canziani; Gomez; Gwirc; Lupi; Moltoni; Nassipián; Slawiski; Zaradnik, "Sistema inalámbrico de microestaciones meteorológicas para aplicaciones agropecuarias". Congreso Argentino de Sistemas Embebidos 2013. ISBN 978-987-9374-88-77.

[2] Canziani; Gomez; Lupi; Nassipián; Slawiski; Turconi; Zaradnik, "Plataforma de conexión de Redes Eléctricas Inteligentes a Internet de las Cosas" en el Congreso Argentino de Sistemas Embebidos 2014. ISBN 978-987-45523-27.

[3] Bernis; Turconi; Benacerraf; Dominguez; Lupi; Zaradnik; Rzepa, "Sistema de seguimiento de dosimetría personal". VII congreso de microelectrónica aplicada 2016. ISBN: 978-987-733-068-7.

[4] Lupi; Zaradnik; Turconi; Dominguez, "Sistema de visualización de precios para supermercados". Congreso Argentino de Sistemas Embebidos 2017 (Case 2017), Buenos Aires, Argentina. ISBN: 978-987-46297-3-9.

[5] Sekhar N. Kondepudi (2016) "Shaping smarter and more sustainable cities. Striving for sustainable development goals". Extraída el 09/03/2018 desde: https://www.itu.int/es/publications/Pages/publications.aspx?lang=en&media=electronic&parent=T-TUT-SSCIOT-2016-1.

[6] La Nación (2017). "Instalaron mil sensores en Buenos Aires: para qué serán utilizados". Extraída el 09/03/2018 desde https://www.lanacion.com.ar/2080116-instalaron-mil-sensores-en-buenos-aires-para-que-seran-utilizados.

[7] "La Ciudad cuenta con una Red de Monitoreo de Aire y Ruido" (n.d.). Extraída el 09/03/2018 desde http://www.buenosaires.gob.ar/noticias/la-ciudad-cuenta-con-una-red-de-monitoreo-de-aire-y-ruido

[8] Telit (2019). "Telit IoT Portal: A Cloud-Based Platform as a Service for IoT". Extraída el 09/03/2018 desde https://www.telit.com/m2m-iot-products/iot-platforms/telit-iot-portal/

[9] Telit (2019). "IoT Portal API Reference Guide". Extraída el 09/03/2018 desde http://help.devicewise.com/display/ARG/IoT+Portal+API+Reference+Guide

[10] Telit (2017). "Telit 3G Modules AT Commands Reference Guide" Extraída el 09/03/2018 desde https://y1cj3stn5fbwhv73k0ipk1eg-wpengine.netdna-ssl.com/wp-content/uploads/2017/11/80378ST10091a_Telit_3G_Modules_AT_Commands_Reference_Guide_r12.pdf

Concurrencia en Arduino: Un enfoque basado en una máquina virtual

Ricardo Moran, Gonzalo Zabala, Matías Teragni

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}@uai.edu.ar

RESUMEN

Arduino es actualmente una de las plataformas más populares para robótica educativa debido a su bajo costo y gran cantidad de recursos disponibles en línea. Las librerías de Arduino proporcionan una capa de abstracción sobre los detalles del hardware, lo que permite construir proyectos interesantes aún sin tener experiencia en el tema. Sin embargo, su falta de soporte para la concurrencia dificulta algunos proyectos de robótica educativa. Como solución, hemos propuesto el uso de un lenguaje de programación concurrente soportado por una máquina virtual que se ejecuta en la placa Arduino. En este artículo, describimos la implementación de dicho lenguaje.

Palabras clave: robótica educativa, lenguaje de programación, máquina virtual, Arduino, concurrencia, programación, sincronización.

CONTEXTO

El presente proyecto está 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 brindado por la misma Universidad Abierta Interamericana.
Ricardo Moran percibe además una beca doctoral por parte de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC).

1. INTRODUCCIÓN

Arduino se ha convertido en una de las plataformas más populares para construir proyectos de electrónica, en especial entre hobbistas, artistas, diseñadores y principiantes en el área. Las librerías de software y el ambiente de desarrollo (IDE) de Arduino proveen una capa de abstracción sobre los detalles de hardware que hace posible construir proyectos interesantes sin una completa comprensión de conceptos más avanzados de microcontroladores como interrupciones, registros, y timers entre otros.
Al mismo tiempo, esta capa de abstracción puede ser ignorada para acceder a funcionalidades avanzadas si el programador lo requiere. Estas características hacen que la plataforma Arduino sea útil tanto para principiantes como para expertos.

Pero hay un aspecto en la que el lenguaje propuesto por Arduino carece de una abstracción adecuada, la concurrencia. Para todo proyecto que contenga cierta complejidad la estructura de setup() y loop() no es suficientemente expresiva. Incluso tareas moderadamente complejas requieren cierta ejecución de tareas simultáneas.
Particularmente, la mayoría de los proyectos de robótica educativa requieren la implementación de un dispositivo que realice dos o más tareas en simultáneo. Esto impone una limitación sobre el tipo de proyectos educativos que se pueden llevar a cabo, especialmente si el contenido que se desea

transmitir no está relacionado a robótica o programación en sí mismo.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Propusimos la implementación de un lenguaje de programación concurrente soportado por una máquina virtual que se ejecuta en Arduino. Lo llamamos UziScript y tenemos la expectativa de que se convierta en el lenguaje al cual compilan entornos de programación visual como Physical Etoys, Scratch for Arduino y Ardublock , entre otros.

Dado que su propósito principal es educativo, fue diseñado en base a los siguientes principios:

1. Simplicidad: la máquina virtual debería ser fácil de entender, comprendiendo cómo realiza su trabajo.
2. Abstracción: el lenguaje debe proporcionar funciones de alto nivel que oculten algunos de los detalles relacionados con los conceptos de microcontroladores básicos y avanzados (como timers, interrupciones, concurrencia, pines de entrada/salida, etc.). Estos conceptos pueden introducirse posteriormente a un ritmo adecuado a las capacidades del estudiante.
3. Supervisión: debería ser posible monitorear el estado de la placa mientras esté conectada a la computadora.
4. Autonomía: los programas deben poder ejecutarse sin tener la placa conectada a la computadora.
5. Depuración: las herramientas deben proporcionar mecanismos para la detección de errores y la ejecución paso a paso del código. Sin herramientas de depuración, el proceso de corrección puede ser frustrante para un usuario sin experiencia.

3. RESULTADOS OBTENIDOS/ESPERADOS

Se diseñó e implementó un entorno de desarrollo integrado para robótica educativa cuyo objetivo es solucionar algunos problemas comunes que se observan en las herramientas actualmente disponibles en el mercado, particularmente la falta de soporte para expresar tareas concurrentes.

La solución tecnológica está compuesta por un firmware que se ejecuta en el robot y un conjunto de herramientas de desarrollo que pueden ejecutarse en una computadora o dispositivo móvil. El firmware desarrollado incluye un programa monitor encargado del protocolo de comunicación y una máquina virtual que abstrae los detalles del hardware. Se diseñó un lenguaje específico de dominio para robótica educativa y se implementaron dos interfaces de programación: una textual y otra visual (basada en programación por bloques). Ambas interfaces fueron diseñadas para usarse en simultáneo, proveyendo generación automática de una representación a partir de la otra.

Dado que la mayoría de los Arduinos contienen un solo microcontrolador, estos solo pueden ejecutar un hilo a la vez. Esto implica que todas las tareas definidas en el programa deben compartir un solo procesador. La máquina virtual, además de ejecutar las instrucciones del programa, es responsable de manejar la planificación de tareas. Esta decide qué tareas son ejecutadas, y cuando deben ser interrumpidas preventivamente. La estrategia de planificación es simple, la máquina virtual recorre la lista de tareas, encolando aquellas cuyo tiempo para correr haya sido alcanzado. Luego ejecuta cada instrucción de dichas tareas hasta que se realice una operación bloqueante, interrumpiendo la tarea actual y pasando a la siguiente. Cada tarea almacena entonces su contexto de ejecución (pila, contador de programa y frame pointer) permitiendo luego continuar la ejecución en el punto donde fué interrumpida. Algunas de las operaciones bloqueantes que fuerzan un cambio de contexto son la instrucción “yield”, todas las variantes de delay(), los saltos a instrucciones anteriores al puntero de ejecución actual, escribir el puerto serie cuando el buffer está lleno, leer del serie cuando el buffer está vacío, entre otras.

Poder simplemente expresar tareas concurrentes no es suficiente, dado que todas comparten los mismos recursos (memoria, pines, puerto serial) es necesario poder coordinarlas. Hemos implementado dos librerías externas que proveen mecanismos de sincronización. Ambas están escritas completamente en UziScript, permitiendo a los usuarios acceder a la lógica de sincronización de ser necesario sin la complejidad de enfrentarse a código de bajo nivel.

Este proyecto se encuentra en un estado de desarrollo avanzado y una primera versión fue publicada a principios de este año. Sin embargo, el lenguaje diseñado todavía presenta muchas limitaciones. Por ejemplo, el único tipo de dato soportado es punto flotante de 32 bits. Sería deseable soportar otros tipos de datos como integers, arrays, y strings.
El firmware también podría mejorarse. Quedan por implementar muchas primitivas de Arduino así como dar soporte a interrupciones de hardware. Estamos trabajando en un diseño que hará posible especificar tareas especiales que sólo se ejecutarán cuando ocurra una interrupción.
Finalmente, la performance también puede ser un problema. Nuestras mediciones iniciales muestran que UziScript es entre 20 y 60 veces más lento que el código Arduino nativo. Si bien esto puede ser suficiente para ciertas aplicaciones (sobre todo en el ámbito educativo), creemos que hay mucho lugar para optimizaciones.
Otro potencial problema es el tamaño de los programas. Considerando que muchas placas Arduino tienen una memoria reducida, es deseable que la máquina virtual y los programas del usuario sean lo más compactos posible. El conjunto de instrucciones, si bien fue diseñado para ser compacto, podría optimizarse. Y el compilador actual no realiza ninguna optimización adicional.

A pesar de las limitaciones descritas anteriormente, creemos que los beneficios que presenta la solución propuesta superan sus desventajas, especialmente en el contexto de robótica educativa.

4. 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.

5. BIBLIOGRAFÍA

- Arduino, "Arduino Playground - Structure," [Online]. Available: http://playground.arduino.cc/ArduinoNotebookTraduccion/Structure. [Accessed 23 Julio 2017].
- M. Resnick, "MultiLogo: A Study of Children and Concurrent Programming," Interactive Learning Environments, vol. 1, no. 3, 1990.
- R. Moran, M. Teragni and G. Zabala, "A Concurrent Programming Language for Arduino and Educational Robotics," in XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017), 2017.
- F. Oudert, "fabriceo/SCoop: Simple Cooperative Scheduler for Arduino and Teensy ARM and AVR," 13 January 2013. [Online]. Available: https://github.com/fabriceo/SCoop. [Accessed 23 July 2017].
- A. Wisner, "wizard97/ ArduinoProcessScheduler: An Arduino object oriented process scheduler designed to replace them all," 15 January 2017. [Online]. Available: https://github.com/wizard97/ArduinoProcessScheduler. [Accessed 23 July 2017].
- I. Seidel, "ivanseidel/ArduinoThread: A simple way to run Threads on Arduino," 15 May 2017. [Online]. Available: https://github.com/ivanseidel/ArduinoThread. [Accessed 23 July 2017].
- K. Fessel, "fesselk/everytime: A easy to use library for periodic code execution.," 2

February 2017. [Online]. Available: https://github.com/fesselk/everytime. [Accessed 23 July 2017].
- G. Bob, "HaikuVM: a small JAVA VM for microcontrollers," 2017. [Online]. Available: http://haiku-vm.sourceforge.net/. [Accessed 15 Junio 2017].
- R. Suchocki and S. Kalvala, "Microscheme: Functional programming for the Arduino," in Scheme and Functional Programming Workshop, Washington, D.C., 2014.
- "PyMite - Python Wiki," 2014. [Online]. Available: https://wiki.python.org/moin/PyMite. [Accessed 15 Junio 2017].
- C. L. Jacobsen and M. C. Jadud, "The Transterpreter: A Transputer Interpreter," Communicating Process Architectures 2004, vol. 62, pp. 182-196, 2004.
- 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.
- M. Elizabeth and C. Hull, "Occam-A programming language for multiprocessor systems," Computer Languages, vol. 12, no. 1, pp. 27-37, 1987.
- A. W. Roscoe and C. A. R. Hoare, "The laws of Occam programming," Theoretical Computer Science, vol. 60, no. 2, pp. 177 - 229, 1988 .
- G. R. Andrews and F. B. Schneider, "Concepts and Notations for Concurrent Programming," ACM Computing Surveys (CSUR), vol. 15, no. 1, pp. 3-43, March 1983.
- Grupo de Investigación en Robótica Autónoma del CAETI (GIRA), "Physical Etoys," 2010. [Online]. Available: http://tecnodacta.com.ar/gira/projects/physical-etoys/. [Accessed 15 Junio 2017].
- Citilab, "About S4A," 2015. [Online]. Available: http://s4a.cat/. [Accessed 15 Junio 2017].
- "Ardublock | A Graphical Programming Language for Arduino," [Online]. Available: http://blog.ardublock.com/. [Accessed 15 Junio 2017].
- J. Smith and R. Nair, Virtual Machines: Versatile Platforms for Systems and Processes, San Francisco, CA: Morgan Kaufmann Publishers Inc., 2005.
- T. Luerweg, "Stack based programming paradigms," in Seminar Concepts of Programming Languages (CoPL), 2015.
- D. Ingalls, T. Kaehler, J. Maloney, S. Wallace and A. Kay, "Back to the future: the story of Squeak, a practical Smalltalk written in itself," in Proceedings of the 12th ACM SIGPLAN conference on Object-oriented programming, systems, languages, and applications, 1997.
- A. Goldberg and D. Robson, Smalltalk-80: The Language and its Implementation, Addison-Wesley Longman Publishing Co., 1983.
- Google, "Google for Education > Blockly," [Online]. Available: https://developers.google.com/blockly/. [Accessed 26 4 2018].
- C. A. R. Hoare, "Communicating sequential processes," Communications of the ACM, vol. 21, no. 8, pp. 666-677, 1978.
- J. Meyerson, "The Go Programming Language," IEEE Software, vol. 31, no. 5, 2014.
- N. Togashi and V. Klyuev, "Concurrency in Go and Java: Performance analysis," in 4th IEEE International Conference on Information Science and Technology (ICIST), 2014.

Desarrollo experimental en Industria 4.0: Ampliación y mejoramiento de un sistema de llenado selectivo de envases con tecnología RFID/NFC.

Lupi, Daniel[1,2]; Zaradnik, Ignacio[1]; Dominguez, Facundo[1]; Rodriguez, Carlos[1]; Kumvich, Augusto[1]; Slawiski, Javier[1]; García, Federico[1].
1Laboratorio de Inteligencia Ambiental Departamento de Ingeniería e Investigaciónes Tecnológicas, Universidad Nacional de la Matanza. Buenos Aires, Argentina; 2Fundación Argentina de Nanotecnología. Buenos Aires, Argentina
izaradnik@unlam.edu.ar

RESUMEN

En el presente trabajo se detallan las ampliaciones y mejoras realizadas al sistema de llenado selectivo de envases basado en tecnología RFID/NFC (Identificación por Radio Frecuencia / Comunicación de Campo Cercano) presentado en el marco del WICC2018 [4]. Habiéndose logrado ejemplificar cómo las tecnologías de RFID/NFC se pueden emplear en procesos productivos, tanto para sumar inteligencia al proceso como para mejorar la trazabilidad del producto, se continuó trabajando sobre el sistema mejorando los aspectos mecánicos, incorporando una etapa de etiquetado previa y ampliando el software para permitir la trazabilidad de toda la cadena de suministros. El sistema está formado por un conjunto de elementos electroneumáticos, un controlador lógico programable, PLC por sus siglas en inglés, un módulo de expansión de entradas y salidas, sensores 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

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 años en el contexto del Laboratorio de Inteligencia Ambiental del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de la Matanza [1][2][3][4][5][6]. Este trabajo se realizó con fondos provenientes del Programa de Investigaciones PROINCE y de un subsidio PICTO del ministerio de Ciencia Tecnología e Innovación Productiva.

1. INTRODUCCION

Tal como se plantea en "Utilización de tecnología RFID/NFC para el desarrollo de un sistema de llenado selectivo de envases" [4], los avances en distintas áreas tecnológicas han dado origen a una nueva industria, la industria 4.0. Este nuevo paradigma plantea dos temas principales [7]: las fábricas inteligentes y la manufactura inteligente. El primero de ellos tiene como objetivo estudiar el sistema de producción inteligente y lograr instalaciones de producción distribuidas en red, mientras el segundo está relacionado 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.

Este trabajo involucra aspectos de ambos temas: el llenado selectivo asociado a la fabricación inteligente y la identificación por RFID/NFC asociado 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 será introducido al comienzo de la línea de producción, momento en el cual se detectará por un sensor óptico y se activará la cinta transportadora. El envase avanzará hasta la posición de etiquetado, donde se detectará por un nuevo sensor y se detendrá su movimiento. El sistema de etiquetado cuenta con un conjunto de rodillos que harán girar el envase mientras que la etiqueta es despegada y adherida al mismo. En la figura N°1 se puede apreciar el sistema de etiquetado.

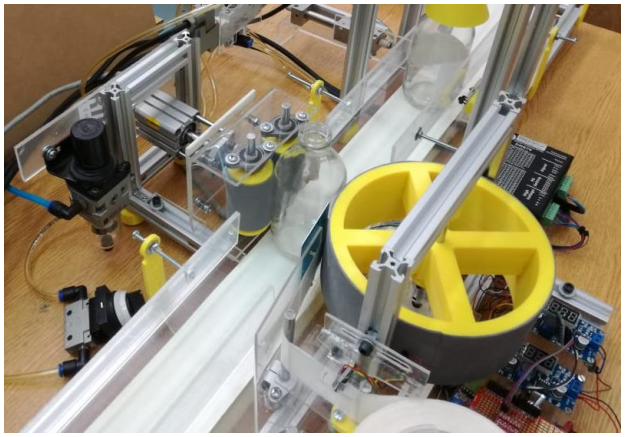


Figura N°1. Sistema de etiquetado.

Finalizado el proceso de etiquetado, la cinta transportadora se activará nuevamente y el envase pasará por un lector NFC, que luego de leer el tag incorporado, lo identificará. En función de la identificación realizada será el proceso de llenado que se produzca, y una vez finalizado este los datos de trazabilidad serán grabados en el tag por otro lector de NFC ubicado al final de la etapa. La figura N°2 presenta una imagen del dispositivo electroneumático encargado de realizar el llenado. El mismo está formado por una cinta transportadora, sensores ópticos para la detección del envase, dos lectores/grabadores de NFC (uno para identificación del envase y otro para la grabación de los datos de trazabilidad del proceso) y varios elementos electroneumáticos para la etapa de etiquetado y llenado del envase.

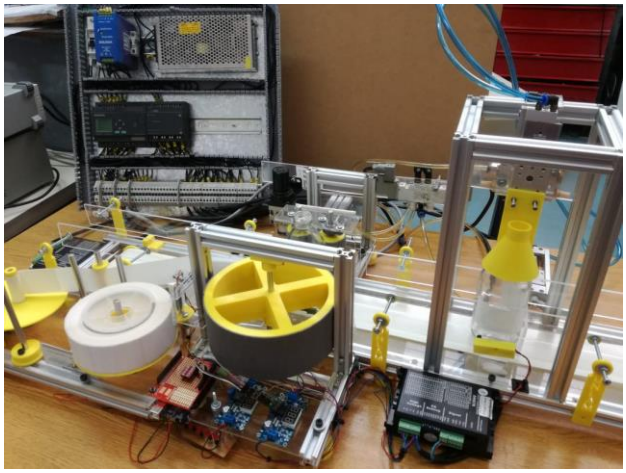


Figura N°2. Dispositivo electroneumático.

El control del dispositivo electroneumático se lleva a cabo por medio de un PLC, o PLR (Relé Lógico Programable), y un módulo de expanción de entradas y salidas, ambos de la firma Siemens. El módulo de expansión se incorporó a raíz del agregado de la etapa de etiquetado. La computadora con la cual cuenta el sistema, a través de un software realizado a medida, se encarga del manejo de los lectores/grabadores de NFC y de indicar al PLC si debe realizar el proceso de llenado o no, lo cual se determina en función de la lectura del tag NFC en el envase. En la figura N°2 se puede apreciar el tablero en donde se encuentra montado el PLC y el módulo de expansión.

3.2. Desarrollo

Si bien al momento de la implementación del sistema se avanzó con las tareas de diseño y fabricación de los tags con electrónica impresa, los cuales se pretenden imprimir en las etiquetas, los resultados obtenidos no fueron satisfactorios. La figura N°3 presenta una imagen de uno de los modelos realizados, mientras que en la figura N°4 se puede apreciar el problema en la continuidad de los trazos presente en las partes curvas de la antena. Como consecuencia de lo expuesto, se continuaron empleando los tags comerciales

basados en el chip NTAG203 de NXP [8]. Si bien hubiera sido conveniente contar con un rollo de los mismos, a fin de usarlos directamente como las etiquetas e integrarlos al proceso, sus costos imposibilitaron su adquisición. Por lo tanto, en el proceso de etiquetado se emplearon etiquetas convencionales, como se puede ver en la figura N°2, y los tags fueron colocados manualmente.

Figura N°3. Tag realizado por electrónica impresa.

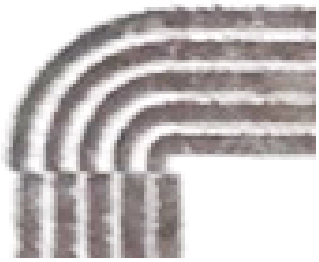


Figura N°4. Problemas de continuidad en trazos curvos.

Los lectores empleados para la aplicación están basados en el chip PN532 de NXP [9], los cuales se comunican con la computadora a través de la interfaz UART.
El nuevo software desarrollado, presenta las siguientes mejoras respecto a su versión anterior:

- En el programa se implementó una estructura de módulos y submódulos similar a la de un sistema ERP, permitiendo a futuro agregar nuevas funciones de una manera fácil y organizada.
- Se incremento el número de tablas de la estructura de bases de datos, de las tres implementadas en la primera versión se llevó a diecinueve en la actual.

La figura N°5 presenta la pantalla principal del nuevo sistema, en donde se pueden apreciar los distintos módulos con que debería contar, habiéndose implementado hasta el momento los módulos de Productos y Producción y encontrándose en desarrollo el de Logística. En la figura N°6 se puede ver la pantalla del módulo de Productos, la cual no cuenta con submódulos y a través de ellas se puede dar de Alta, Baja y Modificar los productos fabricados. La figura N°7 muestra la pantalla asociada al módulo de Producción, en donde se pueden ver los distintos submódulos desarrollados.

Figura N°5. Pantalla principal del Sistema.

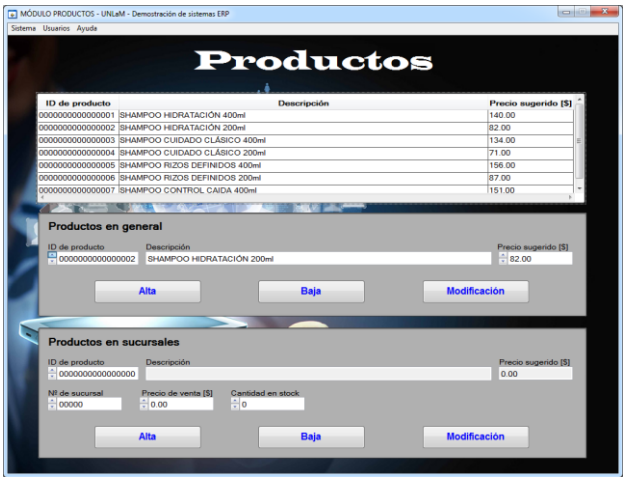


Figura N°6. Módulo Productos.

El submódulo Administrador de producción se encarga de controlar el stock de la materia prima utilizada en la producción. Realiza los pedidos correspondientes en el momento adecuado para mantener un mínimo stock y asegurar un continuo funcionamiento de la línea de producción. Organiza en lotes los productos solicitados en los pedidos recibidos, para maximizar la cantidad de productos iguales

fabricados consecutivamente, consiguiendo así una mayor productividad.

Figura N°7. Módulo Producción.

A través del submódulo Grabación de productos se graban los tags de los envases vacíos con los datos iniciales de producción. El submódulo Identificación de producto se encarga de identificar los productos en el proceso productivo y definir la acción que se debe realizar. El submódulo Grabación/Conteo /Empaquetado se encarga de grabar los datos de producción en los tags de los envases. Este submódulo también realiza un conteo de los productos procesados para, en el posterior proceso de empaquetado, indicarle a la línea de producción cuando una caja debe cerrarse ya sea por estar llena, o por haberse completado la cantidad de unidades de un producto solicitada por un cliente. El submódulo Pallet realiza el control de las cajas que se empaquetan en cada pallet, almacenando esta información en la base de datos para un correcto seguimiento. Finalmente, el submódulo Lector de etiquetas permite leer cualquier tipo de tag, ya sean tags de envases, productos o pallets. Automáticamente se detecta el tipo de tag leído, y se muestra la información almacenada en cada uno de sus campos.

Para el desarrollo de la aplicación de la computadora se empleó la versión de prueba de LabWindows/CVI [10], 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 [11], el cual es un sistema de bases de datos relacional, multihilo y multiusuario con licencia GNU GPL.

Para la programación del PLC se analizaron las distintas entradas y salidas involucradas en el proceso, las cuales aumentaron a raíz de la incorporación del sistema de etiquetado, y el funcionamiento general del dispositivo electroneumático. En base al análisis realizado, se dividió el proceso en pasos secuenciales, donde cada uno de ellos depende de la finalización del paso anterior, y se implemento el programa a través de diagramas de funciones (FUP de la palabra alemana *Funktionsplan)*, programación basada en bloques lógicos del tipo Set-Reset, Timer y AND y OR, ya que al programador le ofrece rapidez y agilidad en dicho proceso.

El aumento de la complejidad del nuevo proceso se refleja claramente en el diagrama de funciones generado. Mientras que la versión anterior, que solo identificaba y llenaba el envase, contaba con cuatro entradas y cuatro salidas discretas, nueve pasos de secuencias utilizando cuarenta y dos bloques de funciones contenido en un diagrama de aproximadamente dos hojas, la nueva versión consta de cinco entradas y ocho salidas discretas, diecinueve pasos de secuencias utilizándose sesenta y tres bloques de funciones representado en un diagrama de cuatro hojas. La figura N°8 y N°9 presentan la primara hoja de ambos diagramas.

3.3. Conclusiones

Al igual que el sistema desarrollado previamente, el nuevo sistema 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.

Las mejoras realizadas, tanto a nivel hardware como a nivel software, han permitido aumentar la flexibilidad del sistema, encuadrándose aún más dentro de los conceptos de Industria 4.0.

Como trabajo futuro está previsto continuar con la ampliación del sistema, incluyendo la etapa de embalaje posterior, que no llegó a implementarse en este periodo. En lo que respecta al software, se plantea finalizar el módulo de Logística.

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 cómo su aplicación puede mejorar los procesos de manufactura. En lo que respecta a la formacion individual, se enumera a continuación: Ignacio Zaradnik la gestión de grupos de trabajos, Javier Slawiski la interpretacion de normas y estandares, Monica Canziani la revison bibligrafica 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 Rodriguez la automatizacion de procesos con RFID y PLC.

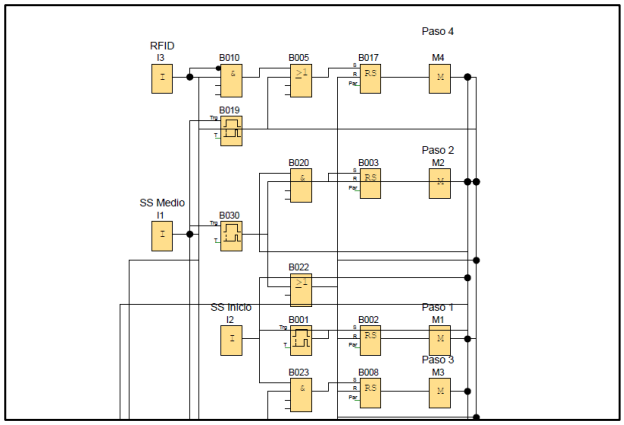


Figura N°8. Diagrama versión anterior.

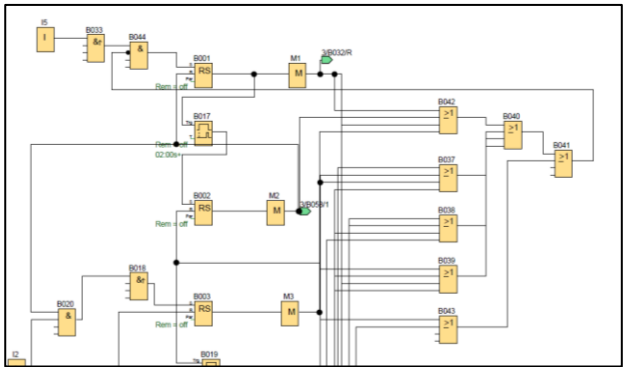


Figura N°9. Diagrama nueva versión.

5. REFERENCIAS

[1] Canziani; Lupi; Ortiz; Slawiski; Zaradnik, "Tecnologías inalámbricas para sistemas de seguimiento, identificación y trazabilidad de productos". uEA 2016, San Luis, Argentina. ISBN 978-987-733-068-7.

[2] Lupi; Zaradnik; Turconi; Dominguez, "Sistema de visualización de precios para supermercados". CASE 2017, Bs.As., Argentina. ISBN: 978-987-46297-3-9.

[3] Canziani; Zaradnik; Tantignone; Lupi; Villares Had, "Procesos Tecnológicos, Tintas y Sustratos empleados en Electrónica Impresa". uEA 2017, Cordoba, Argentina.

[4] Lupi; Zaradnik; Dominguez; Rodriguez; Kumvich; Slawiski; García, "Utilización de tecnología RFID/NFC para el desarrollo de un sistema de llenado selectivo de envases. Industria 4.0". WICC 2018, Corrientes, Argentina.

[5] Lupi; Zaradnik; Canziani; Ortiz; Villares Had; Turconi, "Introducing Printed Electronics in the Electronic Engineering career". EDULearn 2018, Palma de Mallorca, España.

[6] Canziani; Zaradnik; Lupi; Slawiski; Turconi, "Desarrollo y calibración de un banco de testeo para tarjetas de RFID según ISO/IEC 10373-6". uEA 2018, Catamarca, Argentina.

[7] Jiafu Wan; Hu Cai; Keliang Zhou, "Industrie 4.0: Enabling Technologies". 2015 International Conference on Intelligent Computing and Internet of Things (IC1T).

[8] NXP, "NTAG203 NFC Forum Type 2 Tag compliant IC with 144 bytes user", Rev. 3.0 — 17 October 2011.

[9] NXP, "PN532/C1 Near Field Communication (NFC) controller", https://www.nxp.com/docs/en/nxp/data-sheets/PN532_C1.pdf, última visita 09/03/2019.

[10] National Instruments, http://www.ni.com/lwcvi/whatis/esa/, última visita 09/03/2019.

[11] MySQL, https://www.mysql.com/, última visita 09/03/2019.

Diseño y desarrollo de sistema de gestión energética

BIFANO, Lautaro; MACCARONE, Jose

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.
bifanolautaro@frlp.utn.edu.ar; macarone@frlp.utn.edu.ar

RESUMEN

Para que las PyMEs no se vean obligadas a traspasar todos los incrementos energéticos a precio es necesario aplicar herramientas de gestión enfocadas a mejorar el desempeño energético del conjunto, procurando un resultado de menor consumo a igual producción o igual consumo con mayor producción[1]

El objetivo del presente proyecto es diseñar, desarrollar e implementar un sistema en línea para el seguimiento y control en apoyo a las PyMEs con el fin de facilitar un uso más racional de la energía, es decir lograr una mayor eficiencia en el consumo energético de las empresas.

Los datos para el desarrollo se toman de la experiencia adquirida a través de la participación del grupo GyTE (Grupo de Gestión y Tecnología Energética de la UTN FRLP) en análisis energético de procesos productivos y en base a resultados y soluciones aplicadas en empresas PyMEs de la Región.

Palabras clave: Sistema de Gestión Energetico – Software – PyMEs

CONTEXTO

Este trabajo se enmarca en el proyecto "RED TECNOLOGICA NACIONAL SOBRE EFICIENCIA ENERGÉTICA" presentado en el CIDEL Argentina 2014 y en el cual intervienen 8 facultades regionales de la Universidad Tecnológica Nacional. En 2018 el diseño y desarrollo de este proyecto surge del trabajo de integración entre el LINES (Laboratorio de Ingeniería en Sistemas de Información) y el grupo GyTE (Grupo y Tecnología Energética perteneciente los Departamentos de Ingeniería Eléctrica e Industrial) en el cual participa José Maccarone co-autor del proyecto original en 2014. Tanto el LINES como GyTE pertenecen a la Universidad Tecnológica Nacional Facultad Regional La Plata y aportan recursos humanos con perfiles específicos para llevar adelante este proyecto.

1. INTRODUCCIÓN

La energía es un insumo para las empresas. Es un insumo igual para todas ya que se adquiere a un mismo valor y su precio no puede negociarse.
Su sustitución termina siendo generalmente más cara, por lo que únicamente quedan dos caminos, transferirla directamente a los costos del producto o servicio o gestionar su consumo y su uso. En esta última forma es que este insumo se convierte en una variable competitiva, a través de "La Gestión Energética".

Siguiendo con esta línea, el objetivo del presente proyecto es proveer a las organizaciones de los diferentes sectores, de una herramienta capaz de ayudarles en la gestión y la eficiencia energética mejorando así su nivel de competitividad.

El proyecto aborda como base la norma IRAM / ISO 50001 y sus guías de referencia para la Gestión de la Energía en las Organizaciones, siendo las mismas de alcance internacional; contemplando en un futuro los requisitos para poder implementarlas y/o certificarlas.

[1] www.argentina.gob.ar/energia/ahorro-y-eficiencia-energetica - Políticas de estado para el ahorro

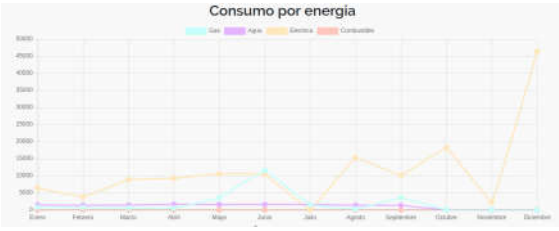
2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Para explicar el desarrollo del Sistema para implementación, seguimiento y control de un Sistema de Gestión Energética (SGE) basado en la norma IRAM / ISO 50001 (en adelante la Norma), se relacionan las acciones reales de campo con el contenido del Software en desarrollo (en adelante SGEfrlp).

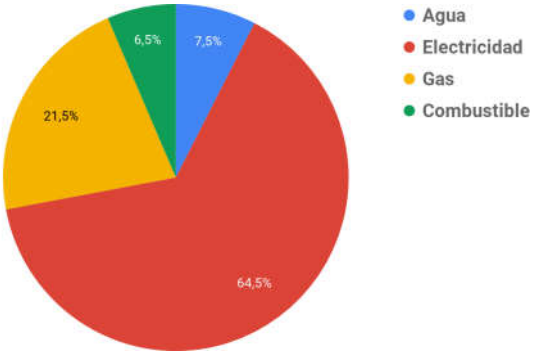
Partiendo de una Revisión Inicial sobre los Consumos de Energía de la PyME se puede obtener una relación de las mismas en sus unidades originales o llevadas a una misma unidad. En este caso serán comparables y se visualizan a través de una Matriz de Consumos Energéticos o a través de un Gráfico tipo Torta.

El SGEfrlp permite que se puedan cargan las facturas de todos los tipos de energía que consume la empresa en sus procesos productivos. De las facturas se especifica el periodo al cual pertenecen, el consumo en la unidad de medida que corresponda según el tipo de energía, el monto que se pagó por dicha energía y datos particulares de cada tipo de energía.

Con la información de todas las facturas tenemos una estadística del consumo energético de la empresa.



Y el gráfico sobre la Matriz Energética, representada en un gráfico tipo Torta, similar al siguiente:



Matriz de Consumos Energéticos

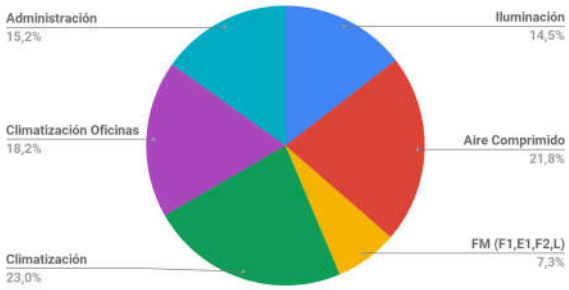
El relevamiento energético de la PyME contiene el estudio de los consumos energéticos de cada proceso productivo. De allí, surge la necesidad de individualizarlos, así también como a cada una de estas partes del proceso involucrado.
Las partes del proceso productivo son denominadas *"Unidades de servicio"* y los productos terminados *"Unidades de producción"*.

Para el estudio de los usos energéticos, denominados en la norma como USEs (usos significativos de energía), se conforman grupos de usos de energía, para lo cual en el SGEfrlp se las llaman *Líneas de Producción* de la empresa. Ellas contienen las *Unidades de servicio* trabajando para obtener *Unidades de producción* o productos.

El SGEfrlp permite que se carguen mediciones del consumo de cada *Unidad de servicio*.

En el caso de no contar con mediciones, permite a través de datos tales como los tipos de energía que utiliza (electricidad, gas, agua o combustibles líquidos, otros), potencia nominal, rendimiento, estado de carga, velocidad real, ciclo de funcionamiento y la cantidad de horas que trabaja al año, obtener una energía equivalente por unidad de tiempo

para cada *Línea de Producción*. De esta manera se puede conocer el peso que tiene cada *Línea de Producción*, la cual también se puede mostrar a través de una Matriz de Uso de Energía o mostrarlo a través de un gráfico tipo Torta. En el gráfico que sigue se muestra un ejemplo de Usos de Energía Eléctrica.



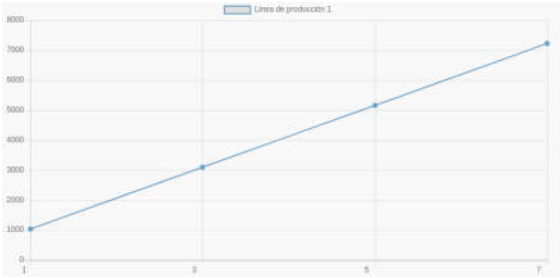
Matriz de Usos de Energía Eléctrica

Un punto de control es dado por la sumatoria de los valores de energía utilizada (suma de usos de energía, llevados a la misma unidad) la cual debe ser aproximadamente igual a la sumatoria de energías consumida (suma de los consumos de energía, llevados a la misma unidad). Si este análisis se realiza mensualmente la matriz de consumos es la suma de las cantidades que figuran en las facturas de electricidad, gas, agua, combustible, otros y será aproximadamente igual a la Matriz de Usos de Energía Total de un mes de producción de la PyME.

De las *Unidades de producción* se obtiene la cantidad de unidades que se producen por hora en cada *Línea de producción*, y sabiendo cuántas horas anuales trabaja dicha *Línea* sabemos cuántas *Unidades de producción* por periodo de tiempo se obtienen.

Uniendo la información de las *Unidades de servicio*, con la proveniente de las *Unidades de producción*, se obtiene el Indicador de cuánta energía de cada tipo contiene una *Unidad de producción* o producto. También es factible obtener la proyección de producción de la empresa y por ende la proyección de la energía que se consumirá para dicha producción.

Esta herramienta brinda a la empresa la posibilidad de trasladar rápidamente a los costos de producción el costo de las energías consumidas en su realización. Ya que cuenta con información del precio de la energía y de la cantidad de energía requerida para producción en tiempo real.



Sabiendo la cantidad de energía que consume la empresa para obtener una *Unidad de producción* es posible proponer una o varias *Líneas Base de Energía (LBEs)*, fotografía inicial o punto de partida que en conjunto con los *Indicadores de Desempeño Energético (IDEs)* permiten a la PyME realizar un seguimiento y control de los planes de acción tendientes a la mejora del *Desempeño Energético.*

En este punto del proceso la empresa conoce su Consumo Temporal (historial de no menos de 12 meses); conoce como usa cada unidad de energía también en forma temporal (anual) y como impacta en la producción. Por lo tanto, la PyME está en condiciones de proponer objetivos y acciones tendientes a mejorar el Desempeño Energético (mayor eficiencia energética).

El SGEfrlp contempla que los *Planes de Acción* responden a una iniciativa concreta de mejora del desempeño energético, como por ejemplo el reemplazo de los Aires Acondicionados de baja eficiencia por los de tecnología Inverter; o pueden responder a las llamadas *No Conformidades*, así llamados los hallazgos provenientes de Auditorías Internas (realizadas por la propia empresa o un tercero) que indicaría desvíos con respecto a los requerimientos de la Norma.

Para el seguimiento y control de los *Planes de Acción* se cuenta con fecha estimada de realización, y un responsable de ejecución. Si no se pudiera cumplir en tiempo y forma, el SGEfrlp permite emitir una *No Conformidad* con el motivo por el cual no se pudo concretar el objetivo propuesto. Para lo cual el SGEfrlp cuenta con el motor necesario para que a través de datos se pueda obtener información valiosa para la toma de decisión. Además provee la documentación de respaldo necesaria para la planificación, seguimiento y control de un Sistema de Gestión de Energía para una PyME.

Todo este camino en busca del objetivo general de la PyME de obtener una nueva *Línea Base Energética* cuya característica principal sea obtener la misma cantidad de *Unidades de producción*, con menos cantidad de *Energía Consumida*.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se espera que los resultados que se obtengan del presente proyecto permitan clarificar la problemática de la gestión de la energía en las PyMEs, y de esta forma, brindar la posibilidad de generar incrementos reales en su competitividad en un entorno regional, nacional e internacional [1].

En particular, el sistema facilita la toma de decisiones en relación a las formas de proceder para hacer más eficiente el consumo energético en una empresa, promoviendo de esta manera a concientizar a la sociedad del derroche de energía, en favor de la recuperación de la crisis energética en la que se encuentra el país.

Se plantea como trabajo futuro, mejorar el sistema producto de este proyecto, SGEfrlp, nutriendo la base de datos con medidores de energía que soporten tecnologías IOT. Estos medidores accederían al sistema desde internet o desde las redes de las empresas a través de una API, para poder registrar información medida y en tiempo real en las bases de datos con información medida.

De esta manera, es posible saber exactamente cuánta energía consume cada *Unidad de servicio* en cada momento, facilitando así la toma de decisiones sobre usos estacionales de la energía en forma online.

4. FORMACIÓN DE RECURSOS HUMANOS

El proyecto tiene como objetivo la formación de RRHH orientados a la programación de sistemas basados en la web y de recursos con conocimientos avanzados en Energías.

Actualmente, es desarrollado por un equipo de dos alumnos becarios del LINES y un docente coordinador los cuales trabajan inter departamentalmente con docentes y becarios del grupo GyTE de los Departamentos de Ingeniería Eléctrica e Industrial de la UTN La Plata.

5. BIBLIOGRAFÍA

[1] RED TECNOLOGICA NACIONAL SOBRE EFICIENCIA ENERGÉTICA - Presentado en CIDEL Argentina 2014

[2] Propuesta para la implementación de un sistema de gestión integral de la energía en la planta de Colombates S.A. - Universidad Nacional de Colombia Facultad de Ingeniería y Administración Palmira, Colombia 2016

[3] 50001:2018 Sistemas de Gestión de la Energía - Requisitos con orientación para su uso.

[4] Norma UNE 216501:2009 Requisitos de las Auditorías Energéticas. España

[5] Source Book for Energy Auditors 1987 - IEA Energy Conservation - MD Lyberg.

[6] Poder Ejecutivo Nacional (2007). Decreto PEN 140/2007 - PRONUREE Programa Nacional de Uso Racional y eficiente de la Energía. Argentina.

[7] Implementación de un Sistema de Gestión de la Energía (SGE). Enrique Bertrán Sánchez.

[8] La crisis energética de la Argentina: orígenes y perspectivas - Roberto Kozulj

Diseño y desarrollo de sistema de gestión tutorial universitaria

LASAGNA, Valeria; ISTVAN, Romina; CHONG ARIAS, Carlos; BIFANO, Lautaro
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.
{valerial, ristvan, c.chong, bifanolautaro}@frlp.utn.edu.ar

RESUMEN

La deserción es un fenómeno presente en todo sistema educativo, relacionado con los procesos de selección, rendimiento académico y eficiencia propia del sistema en general, es decir, el resultado de la combinación y efecto de distintas variables (Díaz Peralta, 2008). Constituye una problemática global que tiene lugar y afecta a todas las instituciones de Educación Superior.

El presente trabajo plantea el diseño y desarrollo de un sistema informático que atienda a esta problemática general, pero particularizando el mismo a las necesidades propias de la institución.

El sistema propuesto permite identificar el riesgo de deserción de cada estudiante; potenciando la comunicación entre tutores, coordinador tutores, ayudantes y docentes, facilitando las tareas de gestión tutorial. Constituye el primer estudio sistemático de la deserción estudiantil en la UTN FRLP.

Palabras clave: Deserción Estudiantil, Deserción Universitaria, Permanencia Académica, Retención académica.

CONTEXTO

El sistema universitario argentino presenta una alta tasa de matriculación, semejante a la que se observa en varios países europeos. Según estadísticas del Ministerio de Educación en el año 2014, Tasa Bruta de Escolarización Superior (TBES) fue del 75,4 % en la Argentina. Esta tasa es, similar a la de Dinamarca y Holanda y superior a la de Austria, Bélgica, Portugal, Reino Unido y Suecia, según estadísticas publicadas por la UNESCO en el año 2015.

Sin embargo, la alta tasa de escolarización universitaria tiene su contracara con una alta tasa de deserción.

Según estudios realizados por el Centro de Estudios de Educación Argentina (CEA) de la Universidad de Belgrano en el año 2017, Argentina se caracteriza, no sólo en Latinoamérica sino también a nivel mundial, por ser uno de los países con menor graduación en proporción al total de estudiantes universitarios. En el total de las universidades argentinas, de cada 100 ingresantes en 2010 se graduaron 30 en 2016. En Brasil esta proporción llega a 50; en Chile, a 60, y en México, a 61 graduados. En muchos países industrializados, esta proporción supera el 70%. Los registros más altos, según la Unesco, corresponden a Japón con 91 y a Dinamarca con 81 (Falco & Istvan, 2017).

Esto resalta que si bien los jóvenes tienen un alto nivel de acceso a la educación superior argentina, la proporción de graduados es una de las menores a nivel global (García de Fanelli, 2014).

El sistema educativo superior enfrenta diversas dificultades a lo largo del trayecto de formación académico que se manifiesta en los altos valores de deserción y repercute directamente en el número de graduados.

Por este motivo, y para responder a las demandas propias de la institución, la UTN FRLP plantea el diseño y desarrollo de un sistema informático para la detección y

caracterización temprana de posibles desertores dentro del Proyecto de Investigación y Desarrollo (PID): "Estudio Sistemático de Deserción Estudiantil Universitaria (ESDEU)", el cual se encuentra homologado y financiado por la UTN con el código TEINNLP0003786.

1. INTRODUCCIÓN

Con el objetivo de realizar el desarrollo del sistema propuesto se profundiza en las características y funcionalidades de herramientas existentes a nivel mundial. De esta manera, se obtienen puntos de referencia sobre el potencial del nuevo software, enmarcándolo en sistemas de tutorías con impacto en la deserción estudiantil.

En función de lo anterior, es posible optimizar los objetivos del nuevo sistema contextualizando los módulos y fuentes de datos necesarias. Consecuentemente, se identifican los actores con sus roles y perfiles, modelando la interacción con el sistema a través de diagramas de casos de uso, evidenciando así, las funcionalidades que permiten la comunicación entre los actores y el software.

Objetivos

El sistema presenta los siguientes objetivos:

1- Consolidar un registro unificado para el estudio sistemático de la deserción, el cual recopile información de diversos medios: base de datos académicas, encuestas, asistencias y actividades de tutores.

2- Identificar tempranamente a posibles desertores mediante la visualización de alertas tempranas.

3- Registrar el avance de estado y seguimiento de los alumnos potenciales desertores y desertores.

4- Brindar un modelo explicativo de los principales factores causales de la deserción y su impacto en cada comisión, materia, carrera e institución en general.

Fuentes de Datos

El sistema propuesto integra los siguientes módulos y fuentes de datos:

- La Base de Datos Académica contiene información personal y académica de los alumnos. En ella se contempla: materias recursadas, libres, aprobadas, notas de parciales, finales, año de ingreso y toda otra información derivada de los mismos.
- Las Asistencias son registradas a fin de poder identificar tempranamente las ausencias consecutivas a una y/o varias materias por parte de los alumnos.
- Las Encuestas tienen objetivos diferentes en cuanto al tipo de información que intenta recabar, dependiendo de la instancia en la que se encuentren los alumnos encuestados.

Para los ingresantes, intenta obtener información acerca de la situación general, el grado de satisfacción respecto al curso de ingreso, aspectos vocacionales, sociales, de integración y desarraigo; profesores y actividades académicas junto con el nivel de exigencia de actividades; la situación económica de los ingresantes y las posibilidades de continuación de dicho periodo.

Las destinadas a los alumnos intermedios y avanzados, focalizan en la actualización de datos personales, situación laboral y familiar, disponibilidad de tiempo para el estudio y dificultades encontradas.

En el año 2018, la UTN FRLP genera un sistema de indicadores de riesgo de deserción estudiantil, en base al análisis realizado mediante técnicas de Minería de Datos sobre alumnos ingresantes en los años 2013/2017. Esta técnica, basada en la extracción de conocimiento, está principalmente relacionada con el proceso de descubrimiento conocido como "Descubrimiento de Conocimiento en Bases de Datos" (KDD, por sus siglas en inglés), el cual se refiere al proceso no-trivial de descubrir conocimiento e información potencialmente útil dentro de los datos contenidos en algún repositorio de información (Han, 2001). Este modelo es integrado al nuevo

sistema para facilitar la detección temprana de alumnos próximos a desertar.

Módulos del Sistema

El sistema cuenta con los siguientes módulos:

- Módulo Alertas / Predictor de Deserción: este módulo permite identificar el grado de riesgo de cada estudiante, visualizar las alertas por materia/comisión, generar listas de estudiantes según los riesgos e identificar los alumnos que necesiten atención prioritaria. El sistema muestra a través de distintos colores alertas visuales con el nivel de prioridad asignada a cada alumno, en relación a la deserción o posible deserción asociada.
- Módulo de Alumnos: permite a Tutores, Jefe de Tutores, Ayudantes y Profesor, Jefe de Cátedra, Jefe de Departamento y Equipo Interdisciplinario visualizar los datos personales, de contacto y perfil académico (año de ingreso, carrera, año de cursada, historial de cursada, analítico, comisión actual, notas de los parciales del corriente año) del alumno, junto con su categorización respecto de la deserción y el acceso al historial de seguimientos o entrevistas realizadas.

Mantiene el registro de inasistencias para cada alumno / materia y su estado de ausentismo.

Permite al alumno comunicarse con el tutor, solicitar tutor y cambio de tutor y a la vez que visualiza alertas relacionadas con su situación académica, como ser vencimientos de cursadas, regularidad, cantidad de horas restantes para quedar libres, aceptación de cambio de comisión.

- Módulo Tutor: este módulo presenta el listado de estudiantes asignados al tutor con la posibilidad de filtrarlos, resalta las alertas de estudiantes posibles desertores identificando con distintos colores cuáles de sus tutorados se encuentran en una situación de deserción crítica, permite el acceso al perfil del estudiante (datos personales, de contacto y perfil académico) y la carga de contactos de seguimiento.

El tutor puede visualizar el indicador y los factores de deserción ya calculados, con la posibilidad de ordenar descendentemente según su indicador de deserción para priorizar la gestión a los registros más críticos. Asimismo facilita la comunicación con el alumno, con docentes y la carga de asistencias/inasistencias, como así también el cambio de categorización de desertor.

El tutor genera cada vez que lo requiera reportes de la situación del alumno y su actividad tutorial.

- Seguimiento de Alumnos: El Tutor es el encargado de registrar la información del seguimiento tutorial: acciones a realizar y observaciones. Las acciones que va a realizar el tutor junto con el alumno para mejorar su desempeño académico, así como las observaciones sobre aspectos importantes que considere, son registradas en este módulo.

El tutor durante su gestión puede registrar: fecha y hora de inicio y fin del contacto, resultado del contacto (si ha contactado al estudiante), observaciones (con la posibilidad de ver un historial de antiguas observaciones realizadas al estudiante), conclusiones.

- Materia/Comisión/Asistencia: visualiza el listado de alumnos por materia/comisión, a fin de poder evaluar la cantidad de desertores y posibles desertores para una materia / curso, posibilitando la visualización y carga de asistencias e inasistencias, y cambios de comisión.
- Módulo Comunicación / Mensajería Interna: permite a comunicarse tanto a docentes, alumnos y tutores entre sí.
- Módulo de Reportes y Estadísticas: este módulo es de gran importancia para fortalecer los procesos de permanencia académica, debido a que proporciona información en tiempo real. Los reportes son generados en archivos Excel y con filtros avanzados para los usuarios.
- Módulo Administrador: gestiona los accesos y permisos por rol a cada módulo del sistema.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

El objetivo general del proyecto es brindar una herramienta informática que facilite la gestión tutorial y caracterice a los estudiantes aportando información no trivial y de utilidad para la gestión.

De esta manera, utiliza conocimiento, metodologías, tecnologías y herramientas de la Ciencia Informática en beneficio de los procesos educativos; promoviendo así, aportes sustanciales en el área de Tecnología Aplicada a la Educación.

3. RESULTADOS OBTENIDOS / ESPERADOS

El futuro sistema integra información histórica académica de los estudiantes, asistencias / inasistencias, resultados de exámenes parciales y finales, encuestas e indicadores de riesgo de deserción.

A través del análisis de esta información, se espera que el nuevo sistema detecte a posibles desertores, mediante la generación de alarmas tempranas. Así, docentes y tutores podrán visualizar a aquellos alumnos en situación vulnerable antes de que el alumno pierda el contacto con la institución.

Constituye el primer estudio sistemático de la deserción estudiantil en la UTN FRLP.

El sistema se encuentra actualmente en su etapa de codificación y se espera su implementación para el ciclo lectivo 2020.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de desarrollo del presente proyecto trabaja desde el año 2014 en diversas áreas relacionadas con la deserción estudiantil universitaria.

Presenta actualmente la siguiente estructura: director (1), docentes investigadores (2), investigador de apoyo (1), tesistas de postgrado (1), tesistas de grado (2), becarios de investigación (2).

5. BIBLIOGRAFÍA

[1] Díaz Peralta C. (2008), "Modelo Conceptual para Deserción Estudiantil Universitaria Chilena", doi: 10.4067/S0718-07052008000200004, Estudios pedagógicos, 34 (2), 65-86.

[2] Falco Mariana, Istvan Romina, Antonini Sergio; University Desertion: Analysis to 2017 admission course in Information Systems Engineering, JAIIO, 2017.

[3] Falco Mariana, Istvan Romina, Antonini Sergio (2017); “Combatiendo la Deserción: Análisis a los Aspirantes 2017 de Ingeniería en Sistemas de Información”; CLEI 2017 / 46 JAIIO, 04 de septiembre de 2017; UTN Facultad Regional Córdoba.

[4] García de Fanelli, Ana (2015); “La cuestión de la graduación en las universidades nacionales de la Argentina: Indicadores y políticas públicas a comienzos del siglo XXI”, Propuesta Educativa Número 43 – Año 24 – Jun. 2015 – Vol1 – Págs. 17 a 31

[5] Han, J.; Kamber M. Data Mining: Concepts ans Techniques. Morgan Kaufmann Publishers, USA, 2001.

[6] Istvan Romina, Chong Arias Carlos, Antonini Sergio; “Sistema de Indicadores de riesgo de deserción para la UTN-FRLP”, Cytal, 2016.

[7] Tinto V. (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.

ENTORNO WEB DE VISUALIZACIÓN DE INFORMACIÓN METEOROLÓGICA PARA EL USO AGRÍCOLA Y DE GENERACIÓN DE ALERTAS ANTE EVENTOS CLIMÁTICOS.

María Masanet*, Flavio Capraro **, Raúl Klenzi*, Martín Muñoz*, Cristian Suarez*

*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
e-mail: {mimasanet, rauloscarklenzi,}@gmail.com

** Instituto de Automática (INAUT), UNSJ - CONICET
Av. Lib. Gral San Martin 1112 (o), San Juan, ARGENTINA,
Teléfonos: 264-4213303, sitio web: www.inaut.unsj.edu.ar
e-mail: fcapraro@inaut.unsj.edu.ar

RESUMEN

El presente trabajo se aborda el desarrollo y puesta en marcha de un sistema de ordenamiento y monitorización de variables climáticas (datos que generalmente son obtenidos desde estaciones meteorológicas automáticas) que permita la presentación de la información en gráficos temporales y en tablas ordenadas. Ello con el fin de que un productor agropecuario logre tomar decisiones acertadas en el manejo de la agricultura moderna que se desarrolla en la provincia de San Juan. Se requiere que el sistema sea del tipo "responsive" a fin de contar con la capacidad de adaptarse a cualquier dispositivo que utilice navegador web. El sistema cuenta con una serie de algoritmos de procesamiento de información que predicen de ocurrencia de contingencias como heladas y viento Zonda, complementado con un motor de eventos que permita enviar alertas o alarmas a los productores registrados al sistema.

Palabras clave: Agricultura Inteligente, Accesibilidad Web, Visualización, Ciencia de Datos.

CONTEXTO

La propuesta se enmarca dentro de los objetivos del proyecto PIO Nº84 "Telemetría Agrícola" que se lleva delante en dependencias del Instituto de Automática de la Facultad de Ingeniería de la UNSJ planteándose como una propuesta de trabajo final de alumnos pertenecientes a la carrera Licenciatura en Ciencias de la Computación del Departamento de Informática FCEFN UNSJ, ámbito en el cual se lleva adelante también, el proyecto "Visualización y Deep Learning en Ciencia de Datos" en el marco del Laboratorio de Sistemas Inteligentes para Extracción de Conocimiento en Datos Masivos del Instituto de Informática de la Facultad de Exactas Físicas y Naturales de la Universidad Nacional de San Juan –FCEFN– UNSJ– desde donde surgirán las propuestas del análisis y procesamiento de datos que

permitan la predicción de ciertas variables climáticas.

1. INTRODUCCIÓN

Hoy en día, en la provincia de San Juan, no se cuenta con una plataforma o sistema informático en línea (bajo el servicio 24/7/365) que permita a los productores locales conocer el estado actual y/o pasado de la climatología de la región, basando la captura de datos desde estaciones automatizadas y bajo los protocolos de instalación para el uso agrícola. En la estación experimental agropecuaria Pocito, del Instituto de Tecnología Agropecuaria, Estación Experimental Agropecuaria (INTA_EEA) San Juan, se cuenta con una estación meteorológica automatizada, cuyos registros se toman cada 10 minutos, y servirá para proveer de la información a una base de datos de la climatología de la zona de Pocito. La mencionada estación meteorológica tiene la capacidad de generar reportes de manera manual por lo que se requiere de una persona en permanente atención y disponibilidad para los productores que requieren de información climática.

Una problemática que afecta seriamente a los productores es la ocurrencia de heladas tardías y de viento Zonda (vientos que superan los 60Km/h, alta temperatura y bajo nivel de humedad ambiente) en el periodo de primavera, donde por lo general los cultivos de tipo frutales suelen ser afectados en las perdidas de flores o congelamiento de yemas. En este sentido, se considera importante contar con una aplicación informática que permita dar una alerta frente a estas contingencias a fin que el productor pueda llevar adelante mecanismos de combate, como así mismo tener un registro continuo de lo ocurrido en el nivel de temperatura.

Se propone tomar la información que se genera desde una estación meteorológica, almacenar los datos en una base de datos, para que una aplicación web lleve a cabo las tareas de ordenar y procesar la información para posteriormente visualizarla en forma de gráficos históricos (series temporales) y en tablas. Se plantea la investigación y desarrollo de algoritmos de predicción de ocurrencia de heladas y viento Zonda, como así también poder determinar las horas frío acumuladas en un periodo de interés. El usuario podrá configurar una plantilla de eventos y alarmas a fin de recibir por email notificaciones de forma automática. También se podrán extraer los registros en una planilla de datos y automatizar un informe de las variables.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La realización de este trabajo ha sido posible debido al financiamiento del proyecto PIO Nº84, “TELEMETRÍA AGRÍCOLA, una herramienta tecnológica para la gestión eficiente del riego, supervisión de cultivos, y generación de alertas” cofinanciado por CONICET y el gobierno de San Juan y con el apoyo brindado por INTA EEA San Juan.
El proyecto se enmarca en la linea de investigación de Riego de Precisión. El riego de precisión se enfoca en la GESTION del riego a fin de minimizar el uso de agua destinada al riego, uniformizar la productividad dentro de un lote y aprender de la gestión de cada temporada para mejorar las estrategias de riego en las futuras campañas. Para ello se requiere investigar y desarrollar nuevas herramientas tecnológicas que permitan recopilación de información de estaciones remotas y en tiempo real. Para la adecuada toma de decisiones, se requiere también investigar y desarrollar algoritmos computacionales que presenten la información de manera ordenada y adecuada; como así también automatizar el envío de alertas o

alarmas ante problemas en el equipo de riego y contingencias climáticas.
Las nuevas herramientas tecnológicas integran desarrollos de sensores, estaciones de medición automáticas, software de procesamiento de datos, algoritmos para la presentación de la información, plataformas en formato web o aplicaciones móviles para hacer llegar los eventos históricos, alarmas y reportes al productor agrícola.
En lo referente al área de Ciencia de Datos y en el ámbito del Departamento e Instituto de Informática se lleva adelante el proyecto CICITCA_UNSJ "Visualización y Deep Learning en Ciencia de Datos" donde se procesan, analizan y visualizan diferente tipología de datos entre otras, series temporales que es el formato de representación que se utiliza en la presente propuesta y desde donde se intentará realizar tareas de predicción mediante la medición de diferentes variables.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Figura 1: captura de pantalla del sitio web http://telemetria.inaut.unsj.edu.ar/inicio/

Atento al objetivo de visualización de los datos capturados desde las estaciones meteorológicas, la Figura 1 muestra la evolución de la evapotranspiración diaria acumulada (barras verticales verdes) y variación de los niveles de temperatura ambiente (máximo, medio y mínimo). En ambos gráficos se registran lo acontecido en durante los últimos 7 días. En forma de tabla se brinda información de los registros climáticos actualizados cada 10 minutos del día en curso.

Resultados esperados:
Hoy en día el productor agropecuario local toma sus decisiones sobre el manejo de riego y del cultivo según su experiencia de temporadas anteriores o visualizando determinadas respuestas en el crecimiento de los cultivos careciendo de registros sobre lo acontecido tanto en clima como en riego y guiado del cultivo.
En este sentido, se espera desarrollar una aplicación web que integre técnicas en diversas áreas, tales como extracción de información, recuperación de información, filtrado de información, representación del conocimiento, gestión del conocimiento, aprendizaje automático, bases de datos, extracción de datos, interacción humano-computadora y la web semántica. Esta nueva herramienta tecnológica permitirá, fundamentalmente a los productores, contar con el acceso y visualización de información, y que le facilitará la toma de decisiones sobre la gestión del sistema de riego y el guiado de los cultivos fundadas o basadas en toda información registradas en la base de datos y sobre parámetros o indicadores calculados automáticamente por la misma aplicación, para que de manera rápida el productor advierta que deben corregirse o adecuarse los procedimientos; así las decisiones son acertadas y con bajo margen de error.
De este modo, en un trabajo de cooperación interinstitucional se generarán conocimientos y tecnologías de gran utilidad para el sector agrícola.

4. FORMACIÓN DE RECURSOS HUMANOS

La temática planteada en esta presentación es llevada adelante como trabajo final de grado en Licenciatura en Ciencias de la Computación y hasta el momento, los alumnos que la elaboran, trabajan en

dependencias del Instituto de Automática (Unidad de doble dependencia UNSJ-CONICET) en donde los alumnos encuentran un asesoramiento de calidad sobre los aspectos físicos y biológicos de las aplicaciones a realizar. En tanto los aspectos referidos a los conceptos de presentación y desarrollo web como así también lo referido a predicciones de variables en el contexto del Data Science se realizan en dependencias del Departamento de Informática DI_FCEFN y el Laboratorio de Sistemas Inteligentes para la Búsqueda de Conocimiento en Datos Masivos del Instituto de Informática, conformando con ello un marco de trabajo de mucha sinergia y enriquecimiento mutuo.

5. BIBLIOGRAFÍA

- **Web Engineering: A Practitioner's Approach.** Author(s): Roger Pressman, David Lowe**.** Publisher: McGraw-Hill Science/Engineering/Math, Year: 2008. ISBN: 0073523291,9780073523293.
- **Building Responsive Data Visualization for the Web**- Hinderman, Bill. Wiley (2016). Published by John Wiley & Sons, Inc.
- **Smart Sensing Technology for Agriculture and Environmental Monitoring-**(Lecture Notes in Electrical Engineering 146) Christos Goumopoulos (auth.), Subhas C Mukhopadhyay (eds.)- Smart Sensing Technology for Agriculture and Environmental Monitoring-Springer-Verlag Berlin HeSpringer-Verlag Berlin Heidelberg, Year: 2012 ISBN: 978-3-642-27637-8,978-3-642-27638-5.
- **Data Science For Dummies,** Lillian Pierson Published by: John Wiley & Sons, Inc (2015).pdf
- **Visualization and Verbalization of Data** Jörg Blasius and Michael Greenacre. Chapman and Hall_CRC (2014)
- **Guidelines for predicting crop water requirements. FAO Irrigation and Drainage** Doorenbos J. and PruittW.O. (1976). Paper 24, 2nd ed. Rome.
- **Crop Yield Response to Water. FAO Irrigation and drainage** Steduto, P., Hsiao, T., Fereres, E., Raes, D., (2012) paper 66.
- **A novel methodology for the monitoring of the agricultural production process based on wireless sensor networks** Escolar Díaz S., Jesús Carretero Pérez, Alejandro Calderón Mateos, Maria-Cristina Marinescu, Borja Bergua Guerra. (2004). Computers and Electronics in Agriculture 76, 252–265 doi:10.1016/ j.compag 2011.02.004
- **Soil-water and solute movement under precision irrigation: Knowledge gaps for managing sustainable root zones. Irrigation** Raine S. R., Meyer W. S., Rassam D. W., Hutson J. L., Cook F. J. (2007). Sci., 26, pp.91-100. Ed. Springer Verlag
- **Review of precision irrigation technologies and their application. National Centre for Engineering in Agriculture**. Smith R. J., Baillie J. N., McCarthy A. C., Raine S. R., Baillie C. P. (2010). University of Southern Queensland. Toowoomba
- **Supervisory control and data acquisition software for drip irrigation control in olive groves. An experience in an arid region of Argentina** Capraro F., Tosetti S., Vita Serman F., (2014). Acta Hortic. vol. 1057, pp. 423-429. DOI: 10.17660/ActaHortic.2014.1057.53
- **TELEMETRÍA AGRÍCOLA. UN ACERCAMIENTO HACIA LAS NUEVAS TECNOLOGÍAS DISPONIBLES EN RIEGO DE PRECISIÓN.** Flavio Capraro, Santiago Tosetti, Vicente Mut. 10º Congreso Argentino de Agroinformática. 47º Jornadas Argentinas de Informática (47 JAIIO). CABA, Argentina. (2018)

ESTRUCTURA Y HERRAMIENTAS PARA UN CURSO INICIAL SOBRE PROGRAMACION ORIENTADA A OBJETOS

Carlos Lombardi (carlos.lombardi@unq.edu.ar) - UNQ / UNAHUR

Nicolás Passerini (npasserini@gmail.com) - UNQ / UNSAM

RESUMEN

Presentamos un proyecto sobre enseñanza de la Programación Orientada a Objetos (POO), en el que se estudian tanto las posibles estructuraciones de un curso inicial sobre POO, como las herramientas a utilizar en el mismo. Partimos de las dificultades que los docentes asociados al proyecto notamos en el dictado de cursos desde hace más de 10 años, y de la literatura existente sobre el tema.
Se formalizó una estructura de curso en cuatro etapas, y se desarrolló un lenguaje y entorno de programación didácticos que acompañan la secuencia sugerida. El entorno incluye facilidades avanzadas de edición, una consola interactiva, soporte para tests de ejecución automática, generación automática de diagramas, errores de codificación adecuados a estudiantes iniciales que aparecen mientras se tipea, entre otras características.

Estos resultados se utilizan en cursos de 5 universidades nacionales desde 2015. También hemos presentado nuestros resultados en eventos nacionales e internacionales. Actualmente estamos trabajando en distintas direcciones con el objetivo de que otros docentes puedan incorporar los resultados de nuestro trabajo.
Hay un tesista de doctorado, uno de licenciatura, y dos becarios estudiantes colaborando en el proyecto.

Palabras clave: programación orientada a objetos, enseñanza de programación, entornos educativos, lenguajes de programación educativos.

CONTEXTO

Este documento describe un trabajo de investigación aplicada, cuyo campo de estudio es la enseñanza inicial de la POO. Los objetos de estudio son las posibles formas de estructurar un curso, y los entornos de programación de propósito educativo adecuados.

Participan docentes pertenecientes a varias universidades, habiéndose generado varios proyectos de investigación. Actualmente está vigente el proyecto PUNQ EXPTE. 1373/17 "Entornos y herramientas de desarrollo orientados a objetos" de la Universidad Nacional de Quilmes. También estuvieron relacionados con el tema proyectos de las Universidades Tecnológica Nacional y Nacional de San Martín , ya concluidos.

Los resultados de este trabajo en curso son aprovechados para el dictado de carreras de informática en la Universidad Tecnológica Nacional (regionales Buenos Aires y Delta), y las Universidades Nacionales de Quilmes, San Martín, Avellaneda y Hurlingham. En todas ellas se utiliza el entorno/lenguaje Wollok, cuyo desarrollo es parte integral del trabajo.

1. INTRODUCCION

El marco de este trabajo son los cursos iniciales sobre POO de nivel universitario, asumiendo que los estudiantes han cursado previamente una materia de programación.
Varios de los docentes que participan en este proyecto llevan más de 10 años dictando cursos iniciales sobre POO en distintas universidades.

En el aprendizaje de los conceptos y técnicas asociados a la POO, confluyen dificultades notadas en la generalidad de los cursos iniciales

de programación, con otras relacionadas específicamente con la POO.
En la experiencia de los docentes que participan en este proyecto, hemos notado en particular:
- la dificultad en el manejo de expresiones complejas, p.ej. el encadenamiento de mensajes o la inclusión de expresiones en los parámetros de una invocación.
- confusiones en el uso una interfaz de objetos para trabajar con listas o colecciones (uso de map, filter, reduce, etc.).
- muy significativamente, problemas para diferenciar los conceptos de clase e instancia, en particular dificultad para comprender la instanciación.
- falta de capacidad para aplicar el concepto de polimorfismo en la resolución de problemas.

El objetivo general del proyecto es propender al dictado de cursos iniciales de POO en los que las dificultades encontradas se mitiguen en la medida de lo posible, y se traten específicamente cuando no se encuentren estrategias adecuadas de mitigación.

Distinguimos dos objetos de estudio que consideramos igualmente relevantes para el logro de los objetivos del proyecto.

El primero está conformado por el planteo y la organización de un curso. En particular, trabajamos sobre el conjunto de conceptos, ideas y técnicas conviene transmitir y practicar en un curso inicial sobre POO; las etapas en las que conviene estructurar un curso; las posibles secuencias didácticas que se derivan de las estructuración propuesta; y la generación de material de apoyo para docentes y estudiantes.

El segundo es el conjunto de herramientas de programación a utilizar en un curso inicial sobre POO. Trabajamos bajo la hipótesis de la conveniencia de utilizar un entorno integrado de desarrollo (IDE, por las siglas en inglés) y un lenguaje de programación concebidos especialmente para su uso en este marco.
Hemos desarrollado sucesivamente varias propuestas de entornos y/o lenguajes, que se realimentan a partir de la experiencia obtenida en los cursos que están a cargo de docentes relacionados con el proyecto.

Los dos objetos de estudio están estrechamente relacionados. Por caso, en la propuesta de estucturación en cuatro etapas que formalizamos recientemente, cada etapa está relacionada con características y construcciones específicas de un lenguaje de programación de propósito didáctico que hemos definido en el marco del proyecto.

Nuestro trabajo se ha nutrido de varios trabajos presentes en la literatura, relacionados con los objetos de estudio de este proyecto.

Respecto del planteo y la organización de un curso, en [Mey03] y [KQPR03] se proponen enfoques progresivos respecto de las tareas que realiza un estudiante a lo largo del curso. Por otro lado, en [BC04] se sugiere basar un curso sobre POO en la construcción de modelos de complejidad progresivamente creciente, en donde el trabajo con los conceptos básicos se da en espiral de acuerdo al patrón definido en [Ber00]. Finalmente, en [CO16] se propone introducir gradualmente los elementos que forman un lenguaje de programación. Las propuestas elaboradas en este proyecto retoman características de todas estas propuestas.

Respecto de los lenguajes y entornos de programación, mencionemos el debate sobre el uso de lenguajes de programación educativos o industriales, cfr. [ACS02], [PS+07].
Dentro de los primeros, existen varias propuestas, algunas para cursos iniciales de programación [MBS12,Pat81] y otras referidas específicamente a POO [BR+96]. En particular el lenguaje Blue [KR96] tiene varios puntos de contacto con Wollok, lenguaje desarrollado en el marco del proyecto.
También hay varios trabajos específicos sobre entornos de programación educativos para POO. Varios de ellos, en particular Traffic [Mey03],

BlueJ [KQPR03] y Dr.Java [ACS02], comparten características con los desarrollados durante este proyecto, p.ej. la inclusión de tests de evaluación automática, un intérprete interactivo para la experimentación rápida, o el cuidado en los mensajes de error que se presentan al estudiante.

También tenemos en cuenta las motivaciones que llevaron a la definición de lenguajes educativos basados en bloques [CDP00,ML07] o en frames [Köl16,MOM15], en particular la posibilidad de utilizar herramientas gráficas en el aprendizaje de la programación.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Definimos dos líneas de trabajo, relacionadas con cada uno de los objetos de estudio que describimos más arriba.

Por un lado, nos proponemos estudiar y definir estructuras y secuencias pedagógicas para un primer curso de POO, suponiendo estudiantes que realizaron previamente un curso básico de programación. Teniendo en cuenta las dificultades encontradas en cursos dictados en años anteriores, que mencionamos en la introducción, concluimos que debe ponerse especial cuidado en los siguientes aspectos:
- introducción de la noción de polimorfismo en la etapa más temprana posible.
- manejo cuidadoso de los conceptos de clase e instanciación.
- trabajo con expresiones, aumentando la complejidad de las expresiones que se utilizan a medida que avanza el curso.

También nos proponemos estudiar en qué punto de un curso inicial sobre POO, y de qué forma, introducir nociones ligadas al desarrollo de software, que no pertenecen al ámbito estricto de la programación. En particular, estamos interesados en testeos automatizados, y en el uso de repositorios de código.

Finalmente, consideramos como un aspecto relevante en nuestros estudios y definiciones, la articulación del primer curso de POO con los siguientes cursos de programación, y la capacidad de aprovechar efectivamente los conocimientos adquiridos en la continuación del trayecto formativo de cada estudiante, y posteriormente en el ámbito profesional.

En la segunda línea de trabajo, investigamos sobre qué características deberían buscarse en un lenguaje y un entorno de programación, para que faciliten el desarrollo de un curso organizado de acuerdo a las secuencias pedagógicas que proponemos.
En particular, evaluamos la conveniencia de utilizar un lenguaje industrial o uno de propósito educativo, y con qué herramientas y características debería contar un entorno de programación para estudiantes.
Durante el desarrollo de esta investigación se generaron varios entornos de programación con propósito educativo. Sucesivamente, surgieron Object Browser, Loop y Ozono [GL+11], todos ellos basados en el lenguaje Smalltalk.
Algunas limitaciones que comprobamos en el uso de estos entornos en cursos que dictamos, nos llevaron a definir un lenguaje de propósito educativo, y acompañarlo con un entorno integrado de desarrollo, también orientado al uso por estudiantes. Así surgió Wollok [JPT14,NL+17,Wol], un producto que combina lenguaje y entorno, usado desde 2015 en varias universidades. Hasta la fecha, Wollok ha sido utilizado en más de 40 cursos universitarios, con un total aproximado de 1500 estudiantes.

Dado que docentes asociados al proyecto dictan cursos de POO en varias universidades, podemos testear continuamente los resultados que vamos obteniendo en el trabajo de investigación, tanto respecto de secuencias pedagógicas, como de herramientas de programación a utilizarse. Esto genera que los resultados de nuestras investigaciones tienen un impacto concreto en la enseñanza de programación en varias universidades, y

además, una realimentación que resulta indispensable en la continuidad del trabajo de investigación y desarrollo.

3. RESULTADOS OBTENIDOS/ESPERADOS

Respecto del trabajo en organización de un curso, hemos podido formalizar una propuesta de estructuración en cuatro etapas o estadíos, que se corresponden con subconjuntos crecientes del lenguaje Wollok.
Las transiciones, tanto de una etapa a la siguiente como considerando el curso como un todo, son objeto de una atención especial.

Entre los trabajos en curso mencionamos:
- sistematización de la base de enunciados de ejercicios, seleccionando enunciados adecuados para cada etapa, y vinculando cada enunciado con los conceptos que se pretende ejercitar específicamente. El objetivo es generar material de guía que facilite la difusión y utilización de los resultados de este trabajo.
- integración con la plataforma Mumuki, que refuerza la posibilidad de difusión, y permite recolectar fácilmente información estadística acerca de la performance de los estudiantes.
- sistematización del análisis cualitativo de resoluciones de ejercicios de modelado utilizando POO, que permitan extraer conclusiones acerca de fortalezas y debilidades de variantes en la organización de un curso. Se individualizaron más de 15 aspectos que pueden estudiarse, y se está preparando el material necesario para llevar a cabo estudios sobre grandes cantidades de estudiantes.

Respecto de las herramientas a utilizarse, tanto los resultados alcanzados como los desarrollos en curso se refieren al lenguaje y entorno Wollok [JPT14,NL+17], y a Wollok-game, una extensión del mismo que facilita la generación de juegos interactivos.

La base de la definición del lenguaje Wollok es estable desde hace varios años, y se ha utilizado exitosamente en varias universidades. El lenguaje admite su introducción en subconjuntos crecientes, que se corresponden con las etapas de la organización de curso que proponemos. Además, su sintaxis permite una transición sencilla a los lenguajes industriales Java y Javascript.

El entorno Wollok es utilizado desde 2015, en más de 40 cursos. Cuenta con varias funcionalidades deseables para un curso inicial sobre POO, entre las que destacamos
- editor con coloreo de sintaxis, autocompletado e indicación de errores mientras se tipea.
- una consola interactiva REPL (Read-evaluation-print loop) que permite experimentar con las características del lenguaje y los objetos y clases desarrollados por el estudiante.
- mensajes de error trabajados para su comprensión por parte de los estudiantes.
- detección de varios defectos de codificación comunes en estudiantes iniciales.
- inclusión de una herramienta para definir y ejecutar tests de evaluación automática.

En 2018 se incorporó Wollok-game, que se utilizó para desarrollar juegos gráficos interactivos en clase.

Entre los trabajos en curso, mencionamos
- el desarrollo de una versión de Wollok que pueda usarse desde un browser Web, evitando la necesidad de instalación.
- la inclusión de una interfaz simplificada con repositorios de código Git.
- varios agregados a Wollok-game, entre ellos soporte de sonido, mejoramiento de las interfaces de código, y un mejor trabajo respecto de las colisiones entre objetos con representación gráfica.

4. FORMACIÓN DE RECURSOS HUMANOS

Colaboran activamente en este proyecto docentes de las universidades en las que se utiliza Wollok, en distintos roles: recolectando y analizando la información sobre el uso de nuestras propuestas y herramientas en el dictado

de materias, formalizando la propuesta se secuencia pedagógica, o generando material para docentes y estudiantes.

Nicolás Passerini está desarrollando su trabajo de Doctorado de Ciencia y Tecnologia en la UNQ sobre esta temática. La tesis de Licenciatura en Informática en la UNQ de Ronny de Jesús consiste en un prototipo de entorno Wollok que pueda utilizarse desde un navegador Web. El trabajo de integración profesional para la Tecnicatura en Programación Informática UNQ de Lucas Sivestri consiste en el desarrollo de una interfaz gráfica de Git a ser integrada dentro de Wollok. El director de estos trabajos es Carlos Lombardi, quien también es director del proyecto de investigación vigente en la UNQ.
Federico Aloi, quien está cursando estudios de profesorado, trabaja como becario de este proyecto, en la sistematización del material para alumnos y la integración de Wollok con la plataforma Mumuki.

5. BIBLIOGRAFIA

[ACS02] E.Allen, B.Cartwright, B.Stoler: "DrJava: a lightweight pedagogic environment for Java", ACM SIGCSE Bulletin 34, 2002.
[BC04] J.Bennedsen, M.E.Caspersen: "Teaching object-oriented programming - towards teaching a systematic programming process". In ECOOP 2004.
[Ber00] J.Bergin: "Fourteen pedagogical patterns". In EuroPLoP 2000.
[BR+96] J.Bergin, J.Roberts, R.Pattis, M.Stehlik: Karel++: a gentle introduction to the art of object-oriented programming. Wiley, EE.UU, 1996.
[CDP00] S.Cooper, W.Dann, R.Pausch: "Alice: a 3-d tool for introductory programming concepts". Journal of Computing Sciences in Colleges 15, 2000.
[CO16] W.Cazzola, D.M.Olivares, "Gradually learning programming supported by a growable programming language". IEEE Transactions on Emerging Topics in Computing 4(3), 2016.
[FPT14] J.Fernandes, N.Passerini, P.Tesone: "Wollok - relearning how to teach object-oriented programming". In Workshop de Ingeniería de Software y Tecnologías Informáticas, UTN, 2014.
[GL+11] C.Griggio, G.Leiva, G.Polito, G.Decuzzi, N.Passerini: "A programming environment supporting a prototype-based introduction to OOP". In IWST 2011.
[Köl16] M.Köllings, "Lessons from the design of three educational programming environments: Blue, BlueJ and Greenfoot.". International Journal of People-Oriented Programming 4, 2016.
[KQPR03] M.Köllings, B.Quig, A.Patterson, J.Rosenberg: "The BlueJ system and its pedagogy". Computer Science Education, 13(4), 2003.
[KR96] M.Köllings, J.Rosenberg: "Blue - a language for teaching object-oriented programming". In SIGCSE 1996.
[MBS12] P.Martínez López, E.Bonelli, F.Sawady "El nombre verdadero de la programación", 2012.
[Mey03] B.Meyer: "The outside-in method of teaching introductory programming". In PSI 2003.
[ML07] D.Malan, H.Leitner: "Scratch for building computer scientists". SIGCSE Bulletin 39(1), 2007.
[MOM15] J.Monig, Y.Ohshima, J.Maloney: "Blocks at your fingertips: blurring the line between blocks and text in gp", in Blocks and Beyond Workshop, IEEE 2015.
[Pat81] R.Pattis: Karel the Robot: a gentle introduction to the art of programming. Wiley, EE.UU, 1981.
[PL+17] N.Passerini, C.Lombardi, J.Fernandes, P.Tesone, F.Dodino, "Wollok: Language + ide for a gentle and industry-aware introduction to OOP". In LACLO 2017.
[PS+07] A.Pears, S.Seidman, L.Malmi, L.Mannilla, E.Adams, J.Bennedsen, M.Devlin, J.Paterson: "A survey of literature on the teaching of introductory programming". ACM SIGCSE Bulletin 39, 2007.
[Wol] www.wollok.org .

Explotación de datos EEG y parámetros fisiológicos de usuarios interactuando en contextos virtuales

Jorge Ierache[1], Gonzalo Ponce[1], Ramiro Nicolosi[1], Claudio Cervino[2], Edgardo Eszter[3] Iris Sattolo[1], Gabriela Chaperon[1],
[1]ISIER-UM -Facultad de Informática Ciencias de la Comunicación y Técnicas Especiales
[2] Instituto de Fisiología y Neurociencias (IFiNe- Secretaria de CyT, UM)
[3]Facultad de Ingeniería
Universidad de Morón
Cabildo 134, Buenos Aires, Argentina
jierache@unimoron.edu.ar

Resumen

La línea de investigación aplicada se centra en el desarrollo de una arquitectura para la explotación de datos registrados durante la interacción de usuarios en un contexto de ambientes virtuales, registrando patrones EEG por la interface cerebro-máquina y datos fisiológicos de usuarios. Se desarrolla con la participación de cuatro Unidades Académicas (Informática, Cs. de la Salud, Cs. Exactas, Químicas y Naturales, e Ingeniería). Existen diferentes maneras de inferir emociones en los usuarios de computadoras, por ejemplo, a través de detección gestual de rostro, voz, ritmo cardiaco y ECG, registros de EEG, este último a través de *Brain Control Interface* (BCI). En el contexto de la computación afectiva actualmente *s*e experimenta con el control de computadoras, dispositivos. Sin embargo, no se dispone de 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, detectar, medir y analizar dichas emociones o respuestas de comportamiento en el ambiente de actuación.

Palabras clave: Computación Afectiva, Estado biométrico, Entornos virtuales, Interfaces cerebro máquina.

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 se complementa con el proyectó "*Influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales*" el que 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; las mismas 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-UM[1] [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-UM[2] 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].

Líneas de Investigación, Desarrollo e Innovación

En particular se orientan los esfuerzos al desarrollo de una arquitectura que facilite el empleo de técnicas y herramientas para el descubrimiento de patrones de datos procedentes de EEG, capturados con el empleo de interfaces cerebro-maquina, datos fisiológicos (ritmo cardiaco).Se suma además en este proyecto captura de rostro del usuario, a fin de determinar su estado emocional, la respuesta del usuario en el marco de una encuesta preguntando al participante en qué estado emocional se encuentra. La misma hace uso de la metodología ampliamente utilizada para capturar emociones: SAM (por sus siglas en inglés Self-Asessment-Manikin) propuesta por Bradley y Lang [15]; frente a su actuación en un contexto virtual, el cual es también capturado y almacenado en correlación con los datos indicados anteriormente. En la actuación del usuario en el contexto virtual, particularmente se aplicarán casos de experimentación empleando el conjunto de datos recopilado por el Center for the Study of Emotion & Attention de la University of Florida en Gaines ville, FL, Estados Unido. El conjunto contiene una base de datos de imágenes (IAPS) [16], que es utilizada en diversos estudios científicos que analizan emociones y es considerada un estándar de facto. En “Estado Emocional Centrado en Estímulos, Aplicando Interfase Cerebro-Maquina”, Ierache et al. [10] (ISIER-UM) presenta una

introducción de BCI, aplicaciones y los resultados obtenidos en el contexto de empleo de sistemas BCI, donde se muestran los resultados iniciales en el desarrollo de un sistema de captura y la lectura de registros del estado emocional de un individuo con la aplicación de interfase BCI, puntualizando en la excitación y meditación del mismo frente a estímulos conformados por imágenes y audio. En esta oportunidad nos proponemos poner a prueba la hipótesis que es posible realizar el descubrimiento de patrones de datos con el desarrollo de capacidades del PIng/17-03-JI-004, influencias del estado biométrico-emocional de personas interactuando en contextos de entornos virtuales [17]. En materia de explotación de datos de la interface biométrica con especial interés en BCI para relacionar los estados anímicos de los usuarios, a los sistemas de adiestramiento y a otros entornos virtuales de aprendizajes, a fin de asistir a la determinación de los parámetros biométricos más relevantes en los cambios motivacionales y emocionales, como así también representar la alteración de señales bioeléctricas contrastándolas y determinando cambios emocionales y el contexto de actuación del usuario.

En la figura 1 se presenta la arquitectura conceptual de explotación de datos EEG, adquiridos con el BCI y Ritmo cardiaco, captura de imagen de rostro y estado emocional interactuando en contextos virtuales estáticos (imágenes IAPS) y dinámicos (juegos, simuladores), con la potencialidad de integrar ambientes inmersivos de realidad virtual integrando registro de encuesta SAM.

Fig 1. Modelo conceptual de la arquitectura del sistema

Objetivos y Resultados Esperados

Los objetivos planteados en el marco de esta línea de investigación aplicada son:

- Desarrollo de nueva base de datos relacional sobre las bases inicial del PIng/17-03-JI-004
- Empleo de imágenes IAPS.
- Integración de ambientes inmersivos de realidad virtual.
- Captura y registración de ritmo cardiaco
- Estudio de algoritmos, técnicas, herramientas para explotación y descubrimiento de datos en particular para el dominio de aplicación.
- Integración de aplicaciones de determinación de estado emocional en función de imagen del rostro.
- Registración de encuesta Self-Asessment-Manikin (SAM).
- Evaluación de Herramientas Open BCI.

Resultados obtenidos

Se realizaros pruebas con ambientes estáticos de imágenes y ambientes dinámico juegos de computadora, empleando un BCI Nueroskype, [18]., registrando valores de EEG y del nivel de relajación y de excitación del usuario en función de cada imagen o escena. Los resultados de las

pruebas mencionadas fueron demostrados y publicados CACIC 2018 [19].

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados, dos investigadores en formación, dos estudiantes de grado.

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. "Brain machine 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’Altto 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. Bradley, M. M. & Lang, P. J. Measuring

emotion: The self-assessment manikin and the semantic differential. J. Behav. Ther. Exp.Psychiatry 25, 49–59 (1994).
16. Lang, P. J., Bradley, M. M. & Cuthbert, B. N. International Affective icture System (IAPS): Technical Manual and Affective Ratings.NIMH Cent. Study Emot. Atten. 39–58, https://doi.org/10.1027/0269-8803/a000147 (1997).
17. Jorge Ierache, Claudio Cervino, Edgardo Eszter, del XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), Facultad de Ciencias Exactas y Naturales y Agrimensura de la Universidad Nacional del Nordeste, 26 y 27 de abril de 2018, ISBN: 978-987-3619-27-4, pp 785-789
18. http://neurosky.com/biosensors/eeg-sensor/algorithms/ vigente marzo 2019
19. Jorge Ierache, Ramiro Nicolosi, Gonzalo Ponce, Claudio Cervino, Edgardo Eszter Registro emocional de personas interactuando en contextos de entornos virtuales, CACIC 2018. Red UNCI, Tandil. Octubre 2018 (en prensa.)

HERRAMIENTAS INFORMÁTICAS PARA EL DESARROLLO DE SERVICIOS DIGITALES INNOVADORES PARA COMUNIDADES URBANAS Y RURALES EN EL MARCO DE CIUDADES Y REGIONES INTELIGENTES

Héctor Luis Vivas, Nicolás García Martínez, Héctor Hugo Ruiz

Laboratorio de Informática Aplicada (LIA) Universidad Nacional de Río Negro Sede Atlántica, Viedma, Río Negro

lvivas@unrn.edu.ar, ngarciam@unrn.edu.ar, hruiz@unrn.edu.ar

RESUMEN

El objetivo principal del proyecto es investigar sobre herramientas informáticas que permitan el desarrollo rápido de servicios públicos para comunidades urbanas y rurales en el contexto del desarrollo de ciudades y regiones inteligentes. En base a nuevos paradigmas y tecnologías digitales disponibles - como, por ejemplo, computación en la nube, datos abiertos, grandes volúmenes de datos, y aplicaciones móviles - el proyecto se focalizará en el diseño y desarrollo de una caja de herramientas informáticas para la prestación de servicios digitales innovadores; en particular, servicios co-creados, anticipados, participativos, y sensibles e inteligentes al contexto. El resultado esperado es una caja de herramientas para gobiernos municipales y actores del sector privado que contenga: a) especificaciones de componentes y b) de servicios a ser utilizados en tiempo de diseño y de ejecución para el rápido desarrollo y despliegue de servicios digitales innovadores, c) especificaciones de buenas prácticas para la integración de dichos componentes y servicios, c) normas para la utilización segura de las herramientas, y d) herramientas informáticas que permitan la visualización de grandes volúmenes de datos para la toma de decisiones de gobierno y la definición de políticas públicas relacionadas con los servicios en el contexto del desarrollo local y regional.

Palabras ***clave:*** Servicios Digitales, Herramientas Informáticas, Ciudades Inteligentes, Regiones Inteligentes, Gobierno Digital

CONTEXTO

El presente trabajo de investigación se realiza en el Laboratorio de Informática Aplicada (LIA) de la Universidad Nacional de Río Negro, cito en la Ciudad de Viedma, Provincia de Río Negro. El proyecto cuenta con financiación de la Universidad nacional de Río Negro Aprobado como proyecto acreditado trienal por Resolución Rectoral 0709 del 13 de Julio de 2017bajo el código PI 40-C-551

1. INTRODUCCION

Con el fin de buscar herramientas que ayuden a proveer soluciones a los problemas acuciantes que deben resolver, los gobiernos recurren a las tecnologías digitales y desarrollan iniciativas de Gobierno Digital. Con nuevas aplicaciones, el área de Gobierno Digital cambia constantemente para reflejar los esfuerzos en innovación que se realizan en el uso de dichas tecnologías en gobierno, identificándose claramente cuatro etapas de evolución: 1) Digitalización – tendiente a la modernización, mejora de la eficiencia interna y acceso a información de gobierno; 2) Transformación – intenta mejorar procesos internos, estructuras y procedimientos de gobierno; 3) Participación – busca transformar las relaciones entre gobierno, ciudadanos, empresas y otros actores no gubernamentales; y 4) Contextualización – las iniciativas apoyan esfuerzos específicos para el desarrollo, es decir para lograr objetivos específicos de políticas públicas y objetivos de desarrollo sostenible. A esta última etapa se la denomina Gobernanza Electrónica Dirigida por Políticas.

Por su parte, la entrega de servicios públicos digitales también presenta innovaciones. Trabajos relacionados muestran un marco de

referencia explicando cómo los sistemas de entrega de estos servicios proveen progresivamente servicios informacionales, mejorados, transaccionales, e integrados, y gradualmente prueban los límites de la innovación incluyendo servicios que promueven la transparencia, son participativos, anticipatorios, personalizados, sensibles e inteligentes al contexto.

Este proyecto se enmarca en la Gobernanza Electrónica Dirigida por Políticas – en particular, en el desarrollo de una caja de herramientas para el desarrollo y la entrega de servicios digitales innovadores que resuelvan necesidades de comunidades urbanas y rurales en el contexto de ciudades y regiones inteligentes. Ejemplos de problemas a abordar en el desarrollo de servicios digitales urbanos y rurales son: 1) facilitar el acceso a la información para promover la co-creación de servicios públicos; 2) provisión de servicios turísticos sensibles e inteligentes al contexto; 3) facilitar la integración de personas con discapacidades físicas a través de servicios diseñados con interfaces adaptativas; 4) facilitar el acceso a información, por ejemplo de precios y nuevos mercados, y la comunicación entre actores de la cadena de mercado, incluyendo agricultores, transportistas, compradores, comerciantes para la venta de productos rurales; 5) la adopción de prácticas agrícolas mejoradas a partir de soluciones tecnológicas – como por ejemplo, riego computarizado, y control computarizado de la pulverización de fertilizantes y plaguicidas, entre otros.

Según datos de censos nacionales, de 2001 al 2010, la población de la provincia de Río Negro creció un 15.5%, y la gran mayoría de esos habitantes (87.1%) se concentró en las ciudades. A fin de mejorar los asentamientos urbanos de la Provincia, como así también la calidad de vida de sus habitantes es necesario planificar, desarrollar y entregar servicios mejorados e innovadores en áreas críticas, como, por ejemplo, el consumo de energía, el tratamiento de la basura, la preservación de los espacios verdes y azules, el transporte automotor, y la inseguridad. Todos estos servicios forman parte de diferentes dimensiones del desarrollo de ciudades inteligentes. Por otra parte, el desarrollo de servicios digitales rurales está alineado con el Plan Estratégico Hortícola 2016-2026 de la Provincia de Río Negro. A modo de ejemplo, como parte de este plan, se firmó una carta de intención con el Ministerio de Agroindustria Nacional para financiar equipos de riego presurizado permitiendo desarrollar cultivos con mayor productividad y de mejor calidad.

Los gobiernos y la población mundial enfrentan grandes desafíos para el desarrollo, a menudo causando pérdida de equilibrio en relaciones y estructuras de poder sociales y económicas; a modo de ejemplo: 1) altas tasas de urbanización y problemas asociados – se prevé que la población urbana mundial crecerá 63% del 2014 al 2050, actualmente 828 millones de personas viven en barrios pobres y carecen de servicios básicos como agua potable y saneamiento (UNDESA, 2014); 2) la falta de empleo estable – del 84 por ciento de la fuerza laboral mundial, el 75% de los trabajadores están empleados con contratos de corto plazo, en empleos informales, o no remunerados (ILO, 2015); y 3) la inequitativa distribución de la riqueza – el 1% de la población posee la mitad de la riqueza mundial (Credit Suisse Global, 2015).

Inserta en el contexto global, la República Argentina presenta similares desafíos: 1) la población total creció un 10,6% en el periodo 2001-2010, y en la Provincia de Río Negro en particular, un 15.5%, donde el 87.1% de este incremento se dio en la población urbana; 2) la tasa de actividad en el país es del 46%; y 3) la desigualdad de la riqueza también constituye un problema central en la Argentina. Estadísticas del INDEC reportan que el 5.1% de la riqueza se distribuye en el 20% más pobre de la población, mientras que el 20% más rico accede al 44.4%.

Considerando los graves desafíos de desarrollo, los países miembros de las Naciones Unidas han adoptado una agenda de desarrollo sostenible que propone 17 objetivos (ODS) a alcanzar en el año 2030. Los ODS consideran particularmente las mejoras a la calidad de vida en diferentes áreas: vida saludable y bienestar (ODS 3); educación y aprendizaje (ODS 4); género (ODS

5); servicios básicos, como agua, saneamiento y electricidad (ODS 6 y ODS 7); empleo y trabajo (ODS 8); conseguir que las ciudades y que los asentamientos humanos sean inclusivos (ODS11); y promover sociedades pacíficas e inclusivas para el desarrollo sostenible, y crear instituciones eficaces, responsables e inclusivas a todos los niveles (ODS 16).

Paralelamente a los problemas de desarrollo, los avances tecnológicos, particularmente de las tecnologías digitales presentan importantes oportunidades y contribuyen con soluciones para mitigarlos. Por un lado, las oportunidades se deben a la gran penetración de estas tecnologías en la sociedad – mundialmente existen más de 7 billones de suscripciones a teléfonos celulares, y 3.2 billones de personas usan Internet (ITU, 2015); y por otro lado, debido a los nuevos desarrollos – dispositivos móviles, computación en la nube, Internet de las cosas, estándares abiertos para interoperabilidad, etc. La República Argentina también se beneficia de las tecnologías: el 53,3% de la población de 3 años o más en viviendas particulares utiliza la computadora, y 7 de cada 10 argentinos usa Internet. Sin embargo, la adopción generalizada de dichas tecnologías crea también una serie de desafíos sociales, económicos, políticos y otros, como amenazas a la privacidad de las personas, riesgos de fraude digital, aparición de monopolios digitales, etc., que tienen que ser abordados con la combinación de tecnología, políticas y cambio social.

Los gobiernos en distintas partes del mundo recurren a nuevas tecnologías digitales a fin de proveer soluciones a los problemas acuciantes que deben resolver. (Janowski, 2015) afirma que el área de Gobierno Digital cambia constantemente para reflejar los esfuerzos en innovación a través de soluciones tecnológicas que realizan los gobiernos para dar respuesta a las presiones sociales, económicas, políticas y otras, y a las transformaciones que en ellos mismos se producen durante estos procesos. El autor identifica cuatro etapas de evolución en el uso de las TD en gobierno: 1) Digitalización – tiene como objetivo la modernización, la mejora de la eficiencia interna y el acceso a información de gobierno; 2) Transformación – intenta mejorar los procesos internos, las estructuras y los procedimientos del gobierno; 3) Participación – busca transformar las relaciones entre gobierno, ciudadanos, empresas y otros actores no gubernamentales; y 4) Contextualización – las iniciativas de Gobierno Digital apoyan esfuerzos específicos de países, ciudades, comunidades y otras unidades sociales para desarrollarse, es decir para lograr objetivos específicos de políticas públicas y objetivos de desarrollo sostenible. A esta última etapa la denomina Gobernanza Electrónica Dirigida por Políticas.

Acompañando la evolución de Gobierno Digital, la entrega de servicios públicos también presenta innovaciones. En (Bertot, Estevez, & Janowski, 2016a), se introduce un marco de referencia que muestra cómo los sistemas de entrega de servicios públicos, apoyados por las tecnologías digitales, proveen progresivamente servicios informacionales, mejorados, transaccionales, e integrados, y gradualmente prueban los límites de innovación incluyendo servicios que promueven la transparencia, que son participativos, anticipatorios, personalizados, sensibles al contexto, e inteligentes al contexto.

Diversos avances tecnológicos permiten alcanzar el desarrollo de servicios digitales innovadores. Los ejemplos incluyen: las transformaciones en gobierno tendientes a proveer ecosistemas que permiten acceder a recursos e innovaciones a través de la colaboración con otros actores (Janssen & Estevez, 2013) (D, Levine, & Wall, 2008); enfoques de gobierno abierto (Mc Dermott, 2010) (Traunmuller, 2014) y de datos abiertos para la co-creación (Koussouris et al., 2012) (Janssen & Zuiderwijk, 2014) y la innovación en la entrega de servicios públicos (Maccani, Donnellan, & Helfert, 2015); estrategias de innovación abierta (Christos et al., 2013) (Mergel & Desouza, 2013) (Finquelievich, 2014), participación ciudadana electrónica (Panopoulou, Tambouris, & Tarabanis, 2010) (Macintosh, Coleman, & Schneeberger, 2009); y el diseño de servicios inteligentes en contextos urbanos (Krassimira et.al., 2011)(Anttiroik, et.al. 2014) y en contextos rurales (Zhang & Liu,

2016)(Esmaeili & Hashemi G, 2015)(Apostol, Leordeanu, Mocanu, & Cristea, 2015).

Por último, existen trabajos relacionados en la adopción de nuevas tecnologías por parte de los gobiernos; en particular, se referencian los últimos trabajos relacionados a soluciones basadas en computación en la nube para gobierno (Wahsh, M.A., Dhillon, 2015) (Bediroglu, S., Yildirim, V., Nisanci, 2016) (Quarati, A., Clematis, A., D'Agostino, 2016) (Dermentzi, E., Tambouris, E., Tarabanis, 2016); a temas de grandes volúmenes de datos para el diseño de servicios públicos (Tayab, M., Zhou, W., Zhao, M., Li, 2016) (Saxena, S., Kumar Sharma, 2016) (Yahaya, J.H., Deraman, A., Abai, N.H.Z., Mansor, Z., Jusoh, 2016) (Lopez et al., 2015) y a temas de datos abiertos para el servicio público (Amugongo, L.M., Nggada, S.N., Sieck, 2016) (Nafis, F., Yousfi, S., Chiadmi, 2016) (Guglielmi, 2016).

2. LINEAS DE INVESTIGACION

Debido a la relevancia de los entornos urbanos en el desarrollo sostenible, uno de los objetivos adoptados por los Países Miembros de las Naciones Unidas a cumplir para el año 2030, en particular el ODS #11, requiere "lograr que las ciudades y los asentamientos humanos sean inclusivos, seguros, resilientes y sostenibles". En este contexto, los gobiernos locales (municipales) dedican importantes esfuerzos para mejorar e innovar en la entrega de servicios públicos para las comunidades urbanas, con el objetivo de mejorar la calidad de vida de los residentes y promover el desarrollo socio-económico, protegiendo los recursos naturales. Del mismo modo, buscan nuevas soluciones para prestar mejores servicios a comunidades rurales, a fin de mejorarles la calidad de vida y evitar que migren a áreas urbanas cercanas. En estos esfuerzos, las tecnologías digitales y los desarrollos informáticos juegan un papel relevante como facilitador e instrumento de innovación en el diseño y desarrollo de estas soluciones. Sin embargo, la mayoría de los gobiernos locales carecen de los recursos humanos, tecnológicos y financieros que les permitan proveer las soluciones que sus comunidades necesitan. En base al problema identificado, el objetivo general del proyecto es investigar el uso de tecnologías digitales para el desarrollo, implementación y entrega de servicios digitales innovadores – es decir, servicios digitales co-creados con los beneficiarios, participativos, anticipados, sensibles e inteligentes al contexto – que sirvan para satisfacer las necesidades de comunidades urbanas y rurales en el contexto del desarrollo de ciudades y regiones inteligentes.

3. RESULTADOS ESPERADOS

Encuadrado en este marco, el proyecto busca desarrollar una caja de herramientas informáticas que permita a los gobiernos locales y a representantes del sector privado innovar en la prestación de servicios. En base a esta premisa, se definen los siguientes objetivos específicos:

O1) Investigar el estado del arte en herramientas informáticas para la entrega de servicios digitales

O2) Proponer metodologías para el diseño y uso de herramientas informáticas para la entrega de servicios innovadores que satisfagan las necesidades de comunidades urbanas y rurales en el contexto del desarrollo de ciudades y regiones inteligentes en la Provincia de Rio Negro

O3) Elaborar un enfoque para la generalización de las herramientas propuestas, en base a estudios comparativos del uso de las herramientas propuestas por O2) y el estado del arte definido por O1)

A fin de alcanzar los objetivos definidos, la hipótesis de trabajo incluye que el proyecto comience con una investigación sobre el estado del arte en el diseño y desarrollo de servicios públicos digitales innovadoras en base a publicaciones científicas, casos de estudio implementados por gobiernos en distintas partes del mundo, y recomendaciones de políticas emitidas por el gobierno nacional, provincial, local, y organismos internacionales. Del mismo modo, se realizarán estudios sobre las necesidades del contexto local. En base a los resultados, se desarrollarán herramientas adecuadas para el desarrollo de servicios específicos, teniendo en cuenta las últimas tendencias identificadas en el estado del arte y las

características del entorno local. Estudios comparativos permitirán evaluar las herramientas desarrolladas, la generalización de resultados, y la definición de enfoques para su transferencia a otros contextos. Durante todo el proyecto, se realizarán esfuerzos para diseminar los resultados obtenidos a través de distintos tipos de publicaciones.

4. FORMACION 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, dos tesis de Maestría, así como para tesinas de grado.

5. REFERENCIAS BIBLIOGRAFICAS

Alarabiat, A., Sa Soares, D., & Estevez, E. (2016). Electronic Participation with a Special Reference to Social Media - A Literature Review. In *Electronic Participation* (pp. 41–52). Springer.

Amugongo, L.M., Nggada, S.N., Sieck, J. (2016). Leveraging on open data to solve city challenges: A case study of Windhoek municipality. In *2016 3rd MEC International Conference on Big Data and Smart City, ICBDSC 2016* (pp. 129–134).

Anttiroik, A.-V., Valkama, P., & Bailey, S. J. (2014). Smart Cities in the New Service Economy: Building Platforms for Smart Services. In *AI & SOCIETY2* (pp. 323–334).

Apostol, E., Leordeanu, C., Mocanu, M., & Cristea, V. (2015). Towards a hybrid local-cloud framework for smart farms. In *2015 20th International Conference on Control Systems and Computer Science, CSCS 2015* (pp. 820–824).

Bertot, J. C., Estevez, E., & Janowski, T. (2016a). Digital Public Service Innovation: Framework, Cases, Trends. In *World Public Sector Report 2016: Innovating Public Service Delivery for Sustainable Developmenty*. United Nations Department of Economic and Social Affairs.

Bertot, J. C., Estevez, E., & Janowski, T. (2016b). Universal and contextualized public services : Digital public service innovation framework. *Government Information Quarterly*, *33*, 211–222. https://doi.org/10.1016/j.giq.2016.05.004

Christos, G., Xenia, Z., Antonis, R., K, P., Jain, A., Gangadharan, G. R., & Yehiia, T. (2013). Applying Open Innovation Strategies to eGovernment for Better Public Services. In *E-Government Success Factors and Measures: Theories, Concepts, and Methodologies* (pp. 308–330). IGI Global.

Dermentzi, E., Tambouris, E., Tarabanis, K. (2016). Cloud computing in eGovernment: Proposing a conceptual stage model. *International Journal of Electronic Government Research*, *12*(1), 50–68.

Douwe, V., Estevez, E., & Janowski, T. (2009). Software Infrastructure for e-Government – e-Appointment Service. In *U- and E-Service, Science and Technology* (pp. 141–152). Springer.

Esmaeili, L., & Hashemi G, S. A. (2015). Rural intelligent public transportation system design: Applying the design for re-engineering of transportation eCommerce system in Iran. *International Journal of Information Technologies and Systems Approach*, *8*(1), 1–27.

Estevez, E., & Janowski, T. (2013). Electronic Governance for Sustainable Development — Conceptual framework and state of research. *Government Information Quarterly*, *30*, S94–S109. https://doi.org/10.1016/j.giq.2012.11.001

Informática y Tecnologías Emergentes

Russo C.[1,2], Sarobe M.[1], Alonso N.[1], Moretti N.[3], Beloso J[3],Ahmad Tamara[1], Gonzalez Carina[4], Serafino Sandra[1], Collazos Cesar[4], Carla Decoud[4]

Instituto de Investigación y Transferencia de Tecnología (ITT)[5]
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento Nro. 1119 3er Piso, Junín (B) – TE: (0236) 4477050 INT 11610

{claudia.russo, monica.sarobe, nicolas.moretti, juanpablo.belos, tamara.ahmad, serafino.sandra}@itt.unnoba.edu.ar

cjgonza@ull.es, ccollazo@unicauca.edu.co, carladecoud@gmail.com

Resumen

Nuestro entorno se va transformando por cambios que se producen con el avance de las tecnologías con un fuerte impacto en nuestra vida cotidiana. En este contexto ¿Qué son las Tecnologías Emergentes?

"Tecnologías emergentes (TE) 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]

Por lo tanto decimos que una TE es el producto que se obtiene al renovar la tecnología que ya antes se ha desarrollado con el fin de obtener mayores beneficios, esto está acompañado de un elevado margen de incertidumbre.

La clave para utilizar al máximo las TE será saber cómo integrarlas de manera efectiva en los productos, procesos o servicios que impactan en nuestra calidad de vida para hacerlos cada vez más personalizados y eficientes.

Este Proyecto de I+D+i es continuación de proyectos anteriores. Se propone identificar, contextualizar, evaluar, desarrollar y aplicar herramientas informáticas usando TE, las cuales tendrán un impacto en contextos heterogéneos.

---

1 Docente Investigador ITT - Investigador Asociado Adjunto sin director CIC

2 Docente Investigador - ITT

3 Becario PROMINF - ITT

4 Docente Investigador externo al ITT

5 ITT - Centro Asociado CIC

Palabras clave: Tecnologías Emergentes, Calidad de vida, Gestión de la Tecnología

Contexto

En el marco del proyecto de investigación "Informática y Tecnologías Emergentes", con lugar de trabajo en el Instituto de Tecnología y Transferencia (ITT) se continúa trabajando en proponer, identificar, contextualizar, evaluar, desarrollar y aplicar diversas herramientas informáticas en tecnologías emergentes, las cuales tendrán un impacto en forma directa en contextos heterogéneos.

El impacto que el surgimiento de todos estos desarrollos tecnológicos tiene en la sociedad es indiscutible. Es innegable el hecho de que están cambiando nuestra sociedad, nuestras costumbres, la forma de relacionarse y comunicarse, la forma en la que las empresas producen, la forma en la que se educa.

La clave para utilizar al máximo las tecnologías emergente será saber cómo integrarlas de manera efectiva en los productos, procesos o servicios que impactan en nuestra sociedad para hacerlos cada vez más personalizados y eficientes.

1. Introducción

Las tecnologías se presentan cada vez más como una necesidad en el contexto de la sociedad donde los rápidos cambios, el aumento de la demanda y actualización de la información y de nuevos conocimientos se convierten en una exigencia permanente.

Es necesario tener presente las transformaciones que como sociedad se ha enfrentado en las últimas décadas, algunas de las cuales persisten hoy en día. Una de estas transformaciones, que continúa motivando cambios en nuestros hábitos y costumbres, ha tenido lugar gracias a la incorporación de

nuevas y más sofisticadas Tecnologías de la Información y la Comunicación (TIC).

Las TIC son herramientas necesarias en la sociedad de la información en general y en particular son útiles en nuestra vida personal, social y laboral. Ofrecen muchas oportunidades: nuevos empleos, participación más amplia en los procesos políticos, participación ciudadana en la toma de decisiones, acceso fácil a información y comunicación con el mundo más allá del hogar y la comunidad, desarrollo de redes, entre otras; aunque también presentan riesgos, mantenerse al margen de las TIC significa desaprovechar todas las oportunidades que ofrecen.

El potencial de las tecnologías emergentes se amplifica cuando interactúan y se combinan de forma innovadora. El impacto se amplifica aún más cuando las tecnologías se unen en plataformas abiertas y ecosistemas. Estos reducen la inversión y plazo de ejecución necesario para impulsar la próxima ola de innovación en los mercados al permitir que las personas y las tecnologías creen rápidamente mayor innovación sobre las olas anteriores. Esta capacidad de combinar y recombinar las tecnologías y aprovechar las innovaciones existentes es el núcleo de las tecnologías emergentes.

La imaginación es la propulsora en la creación de nuevas tecnologías. Haciendo un ejercicio de prospectiva tecnológica podemos anticipar las innovaciones tecnológicas que deberían producirse en el futuro más o menos cercano en campos tan dispares como la nanotecnología, las tecnologías de la información, la ciencia cognitiva, la robótica, la inteligencia artificial, la biotecnología, la salud, la seguridad, la innovación energética o la ciencia de los materiales. [2] En los próximos años las tecnologías emergentes impactarán en el avance tecnológico y cobrarán importancia lo desarrollos que apliquen dicha tecnología.

Líneas de Investigación y Desarrollo

El Proyecto tiene como finalidad investigar cómo la informática impacta en el desarrollo de tecnologías emergentes de manera de analizar, definir y desarrollar herramientas y estrategias innovadoras que impacten de manera responsable en el desarrollo de la

sociedad. Entre las tecnologías emergentes que abordará el proyecto se encuentran la realidad aumentada[3] [4] , la realidad virtual, telepresencia holográfica, robótica, procesamiento de imágenes, sensores, HCI y computación ubicua entre otras aplicada a diversos contextos.

El objetivo general será abordado a través de los siguientes objetivos específicos:

Gestión de la tecnología: Propiciar la investigación (hardware, software) de soluciones informáticas innovadoras utilizando tecnologías emergentes como estrategia para resolver problemas de contextos heterogéneos.

El impacto de las tecnologías emergentes en la sociedad: Propiciar el desarrollo o modificación de un producto, un servicio o un proceso utilizando tecnologías emergentes, con el objetivo de que su efecto sea una mejora en la calidad de vida.

Fomentar la formación de recursos humanos, promoviendo líneas de investigación que permitan la redacción, exposición y defensa de tesinas de grado, tesis de postgrado, y la realización de prácticas profesionales supervisadas.

Redactar, exponer, y defender Tesis de grado y Postgrado (Magíster y Doctorado) y trabajos de Prácticas Profesionales Supervisadas (PPS). Reforzar los conocimientos y buenas prácticas adquiridas en el marco de los trabajos de investigación anteriores, contribuyendo a la formación de nuevos investigadores en el ámbito de la UNNOBA.

2. Resultados Obtenidos/Esperados

En este contexto y como se mencionó anteriormente este proyecto es continuación de proyectos anteriores por lo que es importante destacar que se firmaron convenios específicos de cooperación con empresas y organizaciones gubernamentales y no gubernamentales; se desarrollaron dos propuestas de tesina de grado y tres prácticas profesionales supervisadas y se realizaron publicaciones en congreso. Se pretende seguir difundiendo y transfiriendo los logros alcanzados mediante la presentación y participación en diferentes revistas especializadas, congresos, jornadas y workshops, todos nacionales/internacionales relacionados con la mejora de los procesos, productos y servicios mediante la innovación.

Se espera también que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su rol de investigadores o partícipes activos en equipos de investigación, transferencia e innovación; [5] fomentando la culminación de sus estudios superiores, promoviendo la redacción, exposición y defensa de Trabajos Finales de Grado y Postgrado, y la realización de Prácticas Profesionales Supervisadas y Tesinas de Grado.

Este proyecto se relaciona con otros proyectos y programas que tienen como lugar de trabajo el ITT.

3. Formación de Recursos Humanos

El equipo de trabajo está compuesto por investigadores formados y en proceso de formación , becarios del sistema de Ciencia y Tecnología, Becarios alumnos, Graduados e investigadores externos.

En el término de dos años, se espera contribuir al inicio y concreción de 2 (dos) Tesinas de Licenciatura en Sistemas, 3 (tres) Prácticas Profesionales Supervisadas de la Ingeniería en Informática, 1 (una) Tesis de Especialización, 4 (cuatro) Tesis de Magíster y 3 (tres) Tesis Doctorales.

4. Bibliografía y Referencias

[1] Extraído de IE http://www.ie.edu/

[2] http://jesusgonzalezfonseca.blogspot.com/2011/05/asi-sera-la-tecnologia-del-futuro-i.html

[3] Cabero, J., y García, F. (coords.) (2016). Realidad aumentada. Tecnología para la formación. Madrid: Síntesis

[4] Martínez, S. y Fernández, B. (2018). Objetos de realidad aumentada: percepciones del alumnado de pedagogía. Píxel-Bit. Revista de Medios y Educación, 53, 1-21.

[5] I. B. Rincón, "Investigación científica y tecnológica como factores de innovación", Entelequia: Revista Interdisciplinar, n.° 14, 2012.

Wanda C. Rodríguez Arocho (2018). Herramientas culturales y transformaciones mentales: Las tecnologías de la información y la comunicación en perspectiva histórico-cultural. VOL. 18 NÚM. 2: (MAYO-AGOSTO) DOI 10.15517/AIE.V18I2.33068. Disponible en: https://revistas.ucr.ac.cr/index.php/aie/article/view/33068

Boude Figueredo, Oscar Rafael, Tecnologías emergentes en la educación: una experiencia de formación de docentes que fomenta el diseño de ambientes de aprendizaje. Educação & Sociedade [en línea] 2013, 34 (Abril-Junio): [Fecha de consulta: 4 de marzo de 2019] Disponible en: ISSN 0101-7330
http://www.ie.edu/ LinkedIn: los 25 perfiles profesionales de mayor demanda. Disponible en: http://noticias.universia.es/empleo/noticia/2014/02/04/1079535/linkedin-25-perfiles-profesionalesdemanda.html

Internet Usage and World Population Statistics are for June 30, 2014. http://www.internetworldstats.com/stats.htm

Debray Régis, Vida y Muerte de la Imagen, Paidós, Barcelona, 1994 (1992).

Moore, Gordon E. (1965), "Cramming more components into integrated circuits". Revista Electronics N° 38

The Richard E. Smalley Institute for Nanoscale Science and Techology: http://cnst.rice.edu/

F. García Córdova y M. R. Trejo, La perspectiva de la investigación tecnológica en educación. México: Limusa, 2012

La programación Reactiva y de Actor en entornos de Internet de las Cosas

Nelson Rodríguez[1], María Murazzo[1], Susana Chávez[1], Tatiana Runco[3], Diego Medel[1], Javier Rosenstein[2]

[1] Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2] Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat
[3] Alumno Avanzado Licenciatura en Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Ignacio de la Roza y Meglioli. C.P. 5402. Rivadavia. San Juan, 0264 4234129

nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, schavez@iinfo.unsj.edu.ar, tatuj@gmail.com, dmedel@iinfo.unsj.edu.ar, rosensteinjavier@uch.edu.ar

Resumen

El surgimiento de Internet de las Cosas, la Computación Móvil y la rápida adopción de Cloud Computing han cambiado el panorama que presenta la computación distribuida. Estos sistemas distribuidos masivos generarán una cantidad excesiva de datos, que saturarán a las redes. Llevar servicios de red, cómputo y otras funcionalidades al extremo medio (fog) o final (mist) permitirá el uso eficiente de estos recursos, para mantener la continuidad fluida desde el Edge al Cloud. Además, IoT presenta características singulares y los desarrollos actualmente se realizan con herramientas y lenguajes de programación convencionales no adecuados para este tipo de sistemas. Estos lenguajes en su mayoría son inadecuados para desarrollar aplicaciones altamente distribuidas, debido a que no soportan mensajes asincrónicos y los errores se pueden propagar, entre otras limitaciones. Por otro lado la programación reactiva y el modelo de programación de actor proveen recursos adecuados para el desarrollo de soluciones para IoT.

La presente propuesta de investigación analizará y evaluará las ventajas de la programación de actor y la programación reactiva, aplicándolas a una variedad de entornos de IoT que favorezcan la continuidad del Edge al Cloud.

Palabras clave: *IoT, Actor Model Programming, Reactive Programming, Edge Computing,*

Contexto

El presente trabajo se encuadra dentro del área de I/D Innovación en Sistemas de Software y se enmarca dentro del proyecto de investigación: Orquestación de servicios para la Continuidad de Edge al Cloud, que ha sido aprobado y está en desarrollo para el período 2018-2019. Asimismo el grupo de investigación viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance desde hace más de 18 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

Internet de las cosas (IoT) se puede definir como un conjunto de diferentes sistemas (por ejemplo, carreteras inteligentes, edificios inteligentes, redes inteligentes) compuesto de componentes heterogéneos pero interactivos (por ejemplo, humanos, automóviles, teléfonos inteligentes, pasarelas, medidores inteligentes), tanto individual como

colectivamente, que proporcionan servicios ciberfísicos innovadores. En resumen, puede describirse como un ecosistema denso, a gran escala, abierto y dinámico de entidades sociotécnicas y aplicaciones [1].

A pesar de comunicarse con redes inalámbricas, presentan características distintivas: no existen estándares consolidados por lo que la variedad de tecnologías es importante, existen varias arquitecturas de software, las velocidades de transferencia en general son bajas y la longitud de las estructuras de datos son menores que las redes IP en general.

En 1999, Kevin Ashton del MIT acuñó el término IoT durante una conferencia en Procter & Gamble [2].

Permite que una gran cantidad de dispositivos cooperen para lograr una tarea común. Cada dispositivo individual es pequeño y ejecuta un núcleo de software limitado. La inteligencia colectiva se obtiene de la colaboración distribuida y la comunicación por Internet. Las soluciones correspondientes de IoT forman grandes sistemas de software distribuidos que presentan requisitos profesionales: escalabilidad, confiabilidad, seguridad, portabilidad y mantenibilidad

En la actualidad, IoT puede incluir productos industriales, comerciales, cotidianos (lavavajillas y termostatos) y redes locales de sensores para vigilar granjas y ciudades [3].

Estas 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 [4]; 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 al Cloud como un servidor centralizado, pero pronto se convertirá en un modelo informático insostenible [5]. Se ha comenzado a comprobar que 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 edge computing [6]. En particular si se realiza en los dispositivos, sensores y actuadores se denomina mist computing, mientras que cuando se desarrolla en el extremo medio (fundamentalmente routers o switchers de red) se denomina fog computing.

La programación para este tipo de sistemas se realiza utilizando lenguajes convencionales, los cuales presentan varias limitaciones para desarrollar aplicaciones para IoT. Los mismos no son de naturaleza asincrónica, no soportan aislación de errores y los mismos se pueden propagar (es común tener errores debido a descargas de baterías o falla de conexión por alcance, entre otras causas). Por lo tanto es conveniente utilizar otro tipo de lenguajes, herramientas o frameworks que permitan desarrollar aplicaciones eficientes para estos entornos.

La programación reactiva y el modelo de programación de actor presentan características que dan respuesta a esta problemática.

La Programación Reactiva (RP) es un paradigma de programación donde un sistema se describe en términos de cambio de valores en el tiempo y la propagación de los cambios [7]. En años recientes, RP ha ganado popularidad en varios campos: programación web (React, Vue, y muchos frameworks web similares), desarrollo de aplicaciones móviles (React Native, RxJava, etc.), redes IoT móviles [8], entre otros.

El modelo de programación de actor, si bien fue definido como un modelo matemático en 1973, no ha llegó a difundirse y pocos lenguajes de programación lo han soportado. Recientemente algunos lenguajes de programación como Erlang (Ericsson), Calvin (Ericsson) o Elixir (José Valim - que se ejecuta sobre la máquina virtual de Erlang) o bibliotecas como Akka (Jonas Boner – está escrito en Scala), hacen hincapié en el modelo de programación de actor.

El mismo es inherentemente distribuido, cada actor no comparte estado con el resto de los actores y puede ser ubicado en diferentes máquinas virtuales en diferentes nodos y tienen además re despliegue en tiempo de ejecución.

Programación Reactiva

El término reactivo se ha utilizado para describir las propiedades de ambos lenguajes de programación. y sistemas. Un lenguaje de programación es reactivo si proporciona las construcciones para programadores para implementar de forma declarativa aplicaciones controladas por eventos. Por otra parte, un sistema es reactivo si puede responder a las entradas de manera oportuna [9].

Su concepción y evolución ha ido ligada a la publicación del Reactive Manifiesto [10], que establecía las bases de los sistemas reactivos. Dicho manifiesto indica las condiciones que debe cumplir para lograr el uso eficiente de los recursos y una oportunidad para que las aplicaciones escalen automáticamente.

Un sistema reactivo cumple lo siguiente:

1. Responsive (Sensible): Reacciona a los usuarios. Una aplicación receptiva satisfará las expectativas del usuario en términos de disponibilidad y comportamiento en tiempo real. La latencia es una de las medidas clave a la hora de evaluar qué tan bien funciona un sistema.

2. Elastic (Scalable): Reacciona a la carga. Para poder interactuar continuamente con su entorno, las aplicaciones reactivas deben poder adaptarse a la carga a la que se enfrentan. Utilizando una mayor capacidad computacional cuando sea necesario y funcionando a través de varios nodos de cómputo dependiendo de la carga.

3. Message Driven (Event-Driven): Reacciona a un evento. Las aplicaciones basadas en la comunicación asíncrona reacciona ante eventos discretos como las solicitudes HTTP sin monopolizar los recursos computacionales mientras espera que ocurra un evento. Este nivel natural de concurrencia produce mejor latencia que las llamadas a métodos síncronos tradicionales. Otra consecuencia de la escritura de programas dirigidos por eventos es que los componentes se acoplan de forma flexible, lo que hace que el software sea mucho más fácil de mantener a largo plazo.

4. Resilient (Resistente): Reacciona a las fallas. Debido a que incluso los sistemas de software más sencillos son propensos a fallar (ya sea relacionado con el software o relacionado con el hardware), las aplicaciones reactivas deben ser resistentes a las fallas para satisfacer la demanda de disponibilidad continua. La capacidad de volver a levantarse en caso de que se encuentre con un problema es posiblemente aún más importante cuando se trata de sistemas escalables y distribuidos, debido a que aumenta la probabilidad de fallas en el hardware o la red.

Por esto, algunos autores afirman que la arquitectura del sistema IoT es considerada la base del paradigma de programación reactiva [11].

Modelo de Programación de Actor

El modelo Actor es una teoría matemática de la computación que trata a los "Actores" como primitivas universales de la computación digital concurrente [12]. El modelo ha sido utilizado tanto como un framework para una comprensión teórica de la concurrencia, y como la base teórica para varias implementaciones prácticas de sistemas concurrentes.

A diferencia de los modelos de computación anteriores, el Modelo de actor se inspiró en las leyes físicas. También estuvo influenciado por los lenguajes de programación Lisp, Simula-67 nd Smalltalk72, así como por fundamentos de redes de Petri, sistemas de capacidad y packet switching. El advenimiento de la concurrencia masiva a través de la computación en la nube y las arquitecturas de computadora multi núcleo ha estimulado el interés en el Modelo de actor.

Dicho modelo de programación fue definido por Carl Hewitt en 1973, y presenta ventajas considerables para el desarrollo de aplicaciones para IoT [13].

En IoT los dispositivos pueden dejar de tener conexión momentáneamente por diversas

razones, por ejemplo la intensidad de la señal, razones climáticas, descarga de baterías (baterías solares) y posiblemente tiempo posterior vuelvan a conectarse, debido a la recarga de la batería solar o al alcance de la señal, si es un dispositivo móvil. Esta conexión y desconexión produce fallos que pueden propagar otros fallos en el resto de la red. En el modelo de programación de actor, cada actor se comunica de forma asincrónica con otro lo cual facilita la comunicación cuando hay conexión - desconexión de los dispositivos. Por otro lado, los actores presentan independencia en la ejecución unos de otros, lo cual en caso de falla no la propagaría al resto de la red. Además son muy livianos, por ejemplo en algunos toolkits como Akka ocupan solamente 600 Bytes, lo cual lo hace muy adecuado para entornos donde los dispositivos tienen poca capacidad de procesamiento y memoria

El modelo de programación de actor permite la concurrencia, está basado en memoria distribuida y puede ser aplicado al edge. [14].

El modelo de programación de actor presenta mínimas diferencias con la programación reactiva.

Los objetos reactivos tienen las siguientes ventajas sobre los actores:

- Los mensajes a métodos no definidos son simplemente puestos en cola.
- El modelo de actor carece de mensajes síncronos.
- La entrega asíncrona de mensajes no preserva el orden

Sin embargo, la sincronización puede ser emulada por pares de mensajes asíncronos, y la mayoría de las implementaciones modernas presentan mejoras en estos los tres puntos.

Líneas de Investigación, Desarrollo e Innovación

El desarrollo de la investigación se centrarán en el análisis de los lenguajes de programación, frameworks y otras herramientas que incorporan el modelo de programación de actor y la programación reactiva, en arquitecturas de IoT que permitan desarrollar aplicaciones en el fog o en los dispositivos (mist), tendientes a lograr la continuidad del Edge al Cloud.

Este análisis implica un cuidadoso estudio de las características de cada uno de ellos, evaluando en cada caso la eficiencia para las arquitecturas IoT, la posibilidad de orquestar recursos y de alcanzar una fluida continuidad al Cloud desde las cosas.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos dieciocho años se trabajó en una variedad de sistemas distribuidos y paralelos. Ya sea en 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, como así también con arquitecturas de menores prestaciones (computación móvil y Edge Computing). Dicha experiencia motivó la elaboración del proyecto de investigación del que forma parte la presente propuesta. El grupo ha realizado publicaciones en el área durante los últimos años: dieciocho trabajos de investigación en diferentes Congresos y Jornadas, se realizaron tres publicaciones en revistas científicas y se transfirieron los resultados mediante seis conferencias en eventos científicos y encuentros de divulgación.

Se han aprobado diecisiete tesinas de grado y un trabajo de especialización y se incorporaron dos becarios de investigación categoría alumno y se encuentran en desarrollo tesinas de grado y de maestría.

Objetivos

El objetivo del grupo de investigación es analizar las ventajas de la programación reactiva y del modelo de programación de actor, investigando distintas arquitecturas y evaluando caso donde la resiliencia y la comunicación asíncrona permitan que los sistemas estén exentos de errores.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los docentes-investigadores de la línea de investigación presentada que figuran en este trabajo, de las universidades de San Juan y Champagnat (Mendoza) y cuatro alumnos de la UNSJ.

Cabe destacar que el proyecto marco de la presente línea de investigación incluye investigadores de UNSL y UNLaR y además se estarían realizando dos tesis doctorales, dos maestrías y varias tesinas de grado.

Recientemente se ha aprobado una tesina de grado sobre Evaluación del Modelo de programación de Actor sobre el entorno de Internet de las Cosas y se espera realizar una tesis de maestría sobre Control Topológico para reducción de interferencia en redes IoT. Se espera iniciar otras tesinas más en el área motivo de la presente propuesta de investigación.

Además se espera aumentar el número de publicaciones y también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o actividades de divulgación y asesoramiento a empresas y organismos del estado.

Referencias

[1] C. Savaglio, G. Fortino, M. Zhou (2016), “Towards interoperable, cognitive and autonomic IoT systems: an agent-based approach, in: Internet of Things (WF-IoT)”, IEEE 3rd World Forum on IEEE, 2016, pp. 58–63.

[2] Hewitt C., Bishop P., Steiger R. (1973). "A Universal Modular Actor Formalism for Artificial Intelligence". IJCAI.

[3] Biron J., Follett J. (2016), “Foundational Elements of an IoT Solution The Edge, The Cloud, and Application Development”, O’Reilly Media, Inc.

[4] 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.

[5] Varghese B., Wang N., Nikolopoulos D., Buyya R., “Feasibility of Fog Computing”, arXiv:1701.05451v1cs.DC

[6] 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.

[7] Bainomugisha E., Carreton A. L., van Cutsem T., Mostinckx, S., de Meuter W. (2013), “A Survey on Reactive Programming”, Journal ACM Computing Surveys (CSUR), Vol 45 N. 4, Aug. 2013, https://doi.org/10.1145/2501654.2501666

[8] Shibanai K., Watanabe T. (2018), “Distributed Functional Reactive Programming on Actor-Based Runtime", AGERE ’18, November 5, 2018, Boston, USA.

[9] Mytera F., Scholliersb C., De Meutera W. (2019),"Distributed Reactive Programming for Reactive Distributed Systems", The Art, Science, and Engineering of Programming, vol. 3, no. 3, article 5; 52 pages.

[10] https://www.reactivemanifesto.org/

[11] Romanov E., Troshina G.(2017), “The IoT-architecture on the principles of reactive programming”, 2017 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON).

[12] Hewitt C. (2015), “Actor Model of Computation: Scalable Robust Information Systems”, Cornell University, https://arxiv.org/abs/1008.1459

[13] Patil S., Nair S., Sruthi , Narkhede, B. E. , Pendam D. V. (2016). A 20/20 vision of Internet of things. IOSR Journal of Business and Management. 18. 76-80. 10.9790/487X-1808027680.

[14] 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

Líneas de Investigación del Laboratorio de Sistemas de Información Avanzados: Dinámica de Tecleo, Computación Afectiva, Extracción de Relaciones Semánticas, Blockchain & Smart Contracts.

Jorge Ierache, Hernán Merlino, Enrique Calot, Juan Rodriguez, German Concilio, Federico Rossi, Sabrina Campa, Ezequiel Aceto, Alejandra Ochoa, Nahuel González.

Laboratorio de Sistemas de Información Avanzados, Departamento Computación
Facultad de Ingeniería, Universidad de Buenos Aires.
Av. Paseo Colón 850 - C1063ACV - Buenos Aires -Argentina
Tel +54 (11) 4343-0893 / 4343-0092.

{jierache,hmerlino,ecalot}@lsia.fi.uba.ar

Resumen

El Laboratorio de Sistemas de Información Avanzados (LSIA) de la Facultad de Ingeniería de la Universidad de Buenos Aires (UBA) cuenta con distintas líneas de investigación: Dinámica de Tecleo, Extracción de relaciones semánticas en español, Computación afectiva, Blockchain y Smart Contracts. En varias de estas líneas se han obtenido aportes relevantes al área mientras que otras están comenzando y mantienen objetivos prometedores.

Palabras clave: *LSIA, Procesamiento del lenguaje Natural, Dinámica de tecleo, Computación afectiva, Blockchain, Smart Contracts.*

Contexto

El Laboratorio de Sistemas de Información Avanzados de la Facultad de Ingeniería UBA cuenta con distintas líneas de investigación [LSIA, 2019]. Fue creado en el 2011 y cuenta con diversas publicaciones; se radicó un proyecto UBACyT en el contexto de cadencia de tecleo, en colaboración con el ISIER UM, para el empleo de Interfaces Cerebro-Máquina (BMI). Recientemente se inició un Proyecto de desarrollo Estratégico (PDE) en el marco de patrones de dinámica de tecleo en el contexto *e-commerce.*

Introducción

En este artículo se enuncian las distintas líneas de investigación y su relación con las materias de grado, fomentando la investigación de alumnos de las carreras de Ingeniería y Licenciatura Informática. Se describen a continuación las distintas líneas de investigación del laboratorio.

Líneas de Investigación, Desarrollo e Innovación

El laboratorio cuenta con las siguientes líneas de investigación:

Keystroke Dynamics. Los patrones neurofisiológicos que vuelven única a una firma manuscrita se pueden observar también en el ritmo de tecleo de un usuario [Joyce & Gupta, 1990]. La técnica que analiza este tipo de patrones se llama Dinámica de Tecleo o Keystroke Dynamics [Calot et al. 2014], [González et al. 2015, 2016]. Desde el laboratorio se mantiene una línea de investigación que, además de articular keystroke dynamics con la extracción de información emocional del usuario, investiga mejores algoritmos de identificación de personas mediante su ritmo de tecleo.

Computación afectiva. Esta línea de investigación se centra en el análisis de factores emocionales de grupos de personas y en la toma de decisiones por software basada en las emociones inferidas. En particular, se realizan estudios sobre grupos de personas en distintos escenarios, detectando sus expresiones, midiendo ritmo cardíaco y monitoreando su actividad cerebral. Con ellos, se permite inferir su estado emocional (empleando métodos de aprendizaje automático supervisado) y se ejecutan acciones que refuerzan o retraen dicho estado.

Extracción de relaciones semánticas en español. Esta línea se orienta al desarrollo de un *parser* semántico, ya sea supervisado o no, que sea capaz de extraer relaciones semánticas de un texto no estructurado: lenguaje natural en español. La salida de este *parser* sería una red semántica, la cual tendría que servir para generar, o bien para extender una ontología de forma semi-automática. Este *parser* estaría compuesto por una serie se pasos, en donde cada uno realizaría una tarea atómica utilizando algoritmos conocidos y probados dentro del área de PLN: Tokenización, POS *tagger*, *named entity recognizer*, *syntactic parser*, entre otros.

Blockchain & Smart Contracts. Esta línea se enfoca en el desarrollo de un framework para la automatización de Smart Contracts (SC) y en todo el proceso de implementación de los mismos en la Blockchain. Para la generación automática de SC se está utilizando programación genética, sistemas basados en conocimiento y sistemas de aprendizaje automático para la evaluación de los mismos.

Resultados y Objetivos

Keystroke Dynamics. Se mostró que el método de Mahalanobis no es siempre útil, especialmente si los datos son linealmente independientes, y que puede ser reemplazado por el método menos costoso de distancia euclídea normalizada [Calot et al., 2013]. A su vez, se creó un conjunto de datos de dinámicas de tecleo no estructurada que fue publicado en el sitio web del laboratorio [Calot, 2015]. Mediante este conjunto de datos y pruebas internas se comenzaron a evaluar estrategias más sofisticadas de extracción de características, filtrado del texto de entrada, selección de modelos y la alternancia de clasificadores que prometen mejorar el EER por debajo del 5% –frente a un 10% en usuarios estándar y 6% en usuarios de mayor pericia dactilográfica– y disminuir su variación entre usuarios. Se encuentra en desarrollo un proyecto de desarrollo estratégico vinculado a la transferencia a la industria reconocido por

la UBA (PDE-44-2019 Reconocimiento de patrones de teclado en ambientes Web). En este marco se realizó un trabajo colaborativo con una de las principales empresas de *e-commerce* de América Latina, empleando técnicas de modelado de cadencia de tecleo para autenticar al usuario, y luego identificar patrones de tipeo compartidos por uno más usuarios de la plataforma. Los resultados experimentales sobre una muestra de 2000 usuarios seleccionados al azar mostraron que es posible identificar a quien tipea con una precisión del 89%, [Concilio et al., 2018], autenticar sesiones con un EER por debajo del 8% y detectar usuarios controlados por una misma persona física, utilizando sólo las marcas de tiempo originadas por la presión de cada una de las teclas del teclado en campos de texto libres. En la muestra se lograron aislar 30 usuarios con comportamiento fraudulento confirmados.

Computación afectiva. Se está recolectando información de los estados de los individuos para nutrir la red de conocimientos de su comportamiento y permitir posteriormente inferir estados. El trabajo realizado consistirá en un framework capaz de detectar el estado predominante del grupo y permitir llevar el mismo a un estado emocional compartido entre todos los miembros del grupo. [Ierache et al., 2012], [Ierache et al., 2014], [Calot, E., Ierache J. et al, 2019]. El empleo de Interfaces Cerebro-Máquina BMI se realizó con el ISIER UM. En este sentido, se está incursionando en el reconocimiento de estados emocionales a partir del registro de bioseñales humanas. Particularmente, se explotará el uso de marcadores biométricos provenientes de señales electroencefalográficas (EEG) y de ritmo cardíaco (HR). En el LSIA se desarrolló una cámara de inmersión emocional para la realización de experimentos [Rossi F. et al., 2017], [Calot E., Rossi F. et al, 2016] [Calot E., Ierache J., 2017]. Esta cámara propició la creación de un ambiente de estimulación con imágenes IAPS (International Affective Picture System), [Lang, P. J., Bradley, M. M. & Cuthbert, B, 1997], audios y videos, en el contexto de múltiples estados afectivos enmarcados en el modelo de Excitación-Valencia [James A. Russell, 1980, 1989]. Durante la estimulación a voluntarios se solicitó el tipeo de textos (en este orden se fusionaron las líneas de investigación de dinámica de tecleo y de computación afectiva). Además, se presentaron encuestas SAM (Self-Assessment Manikin) [Bradley, M. M. & Lang, P. J., 1994] para ser respondidas luego de cada estimulación. Si bien el EEG como base de inferencia ha sido tratado ampliamente, las líneas de investigación proponen establecer un marco neurocientífico sobre el cual se sustenten los criterios de diseño y ejecución experimentales aplicados. A su vez y, dado que el ritmo cardíaco no resulta ser una señal de uso extensivo en estudios de computación afectiva, se evaluará su robustez como característica de inferencia, analizando el desempeño de múltiples algoritmos de aprendizaje automático, brindando así un aporte al estado del arte del área.

Extracción de relaciones semánticas en español. Esto corresponde a una investigación en proceso en la cual ya se han obtenido resultados en un experimento realizado con prescripciones médicas, en donde se probaron dos métodos de clasificación de textos, con textos médicos en español: *Naïve Bayes Multinomial* (NBM) y *Support Vector Machines* (SVMs). La clasificación de textos, en este caso, si bien fue acotada al campo de la medicina, logró una precisión del 86% [Rodríguez et al., 2014]. Otros

avances en esta línea se han publicado en [Rodríguez et al., 2015], [Rodríguez et al., 2016a], [Rodríguez et al., 2016b], [Rodríguez et al., 2017a] y [Rodríguez et al., 2017b]

Blockchain & Smart Contracts. Los resultados hasta ahora obtenidos permiten esperar que el desarrollo de los mismos pueda ser llevado adelante por personas sin experticia en el tema, solo contestando un conjunto de preguntas. Un sistema experto generará el contrato y permitirá su puesta en etapa de pruebas o producción según lo defina el usuario.

Formación de Recursos Humanos

El laboratorio actualmente se conforma de dos investigadores formados, tres investigadores en formación, tres alumnos investigadores, un egresado investigador en formación. Se han radicado dos tesis de doctorado y diversas tesis de grado de la Facultad de Ingeniería de la UBA. Los docentes participan de materias de la carrera de Ingeniería en Informática y Licenciatura en Sistemas: (75.39) "Aplicaciones Informáticas", (75.41) "Algoritmos y Programación II", (75.50) "Introducción a los Sistemas Inteligentes", (75.65) "Manufactura Integrada Por Computador",(75.66)"Manufactura Integrada Por Computador II", (75.67) "Sistemas Automáticos de Diagnóstico y Detección de Fallas I", (75.68) "Sistemas de Soporte para Celdas de Producción Flexible", (75.69) "Sistemas Automáticos de Diagnóstico y Detección de Fallas II" y (75.70) "Sistemas de Programación no Convencional de Robots".

Referencias

BRADLEY, M. M. & LANG, P. J. Measuring, Emotion: The Self-Assessment Manikin and the semantic differential. J. Behav. Ther. Exp.Psychiatry 25, 49–59, 1994

CALOT, E. 2015. "Keystroke Dynamics keypress latency dataset". Base de datos para investigación. http://lsia.fi.uba.ar/pub/papers/kd-dataset/

CALOT, E.; PIRRA, F.; RODRIGUEZ, J.M.; PEREIRA, G.; IRIBARREN, J.; IERACHE, J. 2014. "Métodos Adaptativos de Educción de Dinámica de Tecleo Centrado en el Contexto Emocional de un Individuo aplicando Interfaz Cerebro Computadora". XVI Workshop de Investigadores en Ciencias de la Computación, ISBN 978-950-34-1084-4.

CALOT, E.; RODRIGUEZ, J.M.; IERACHE, J. Improving versatility in keystroke dynamic systems. En Proceedings del XIX Congreso Argentino de Ciencias de la Computación, number 5606, 2013. ISBN 978-987-23963-1-2. URL http://hdl.handle.net/10915/32428

CALOT, E.; RODRIGUEZ, J.M.; IERACHE, J., 2014. Improving versatility in keystroke dynamic systems. En Jorge Raúl Finochietto y Patricia Mabel Pesado, editors, Computer Science & Technology Series. XIX Argentine Congress of Computer Science, Selected papers, páginas 289–298. Editorial de la Universidad Nacional de La Plata (EDULP), 2014b. ISBN 978-987-1985-49-4. URL http://lsia.fi.uba.ar/papers/calot14b.pdf

CALOT, E., ROSSI F., GONZÁLEZ N., HASPERUÉ, W., IERACHE, J.. Avances en educción de dinámica de tecleo y el contexto emocional de un individuo aplicando interfaz cerebro computadora. En WICC 2016, Entre Ríos, Argentina), páginas 872–876, jun. 2016. ISBN 978-950-698-377-2. URLhttp://hdl.handle.net/10915/53247

CALOT, E., IERACHE, J.. Multimodal biometric recording architecture for the exploitation of applications in the context of affective computing. En Proceedings del XXIII Congreso Argentino de Ciencias de la Computación (La Plata, 2017), number 10529, páginas 1030–1039, 2017. ISBN 978-950-34-1539-9. URL http://hdl.handle.net/10915/63866

CALOT, E., IERACHE, J., HASPERUÉ, W.. 2019. Robustness of keystroke dynamics identification algorithms against brain-wave variations associated with emotional variations.

En Advances in Intelligent Systems and Computing. Springer, c. En prensa

CONCILIO, G., IERACHE, J., MERLINO, H., CALOT, E.. Application of Keystroke Dynamics Modelling Techniques to Strengthen the User Identification in the Context of E-commerce. En XXIV CACIC 2018. En prensa. URL http://lsia.fi.uba.ar/papers/concilio18.pdf

ROSSI, F., CALOT, E., IERACHE J. Educción emocional de un individuo en contextos multimodales en computación afectiva. En XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, Buenos Aires, Argentina), jul. 2017. ISBN 978-987-42-5143-5

IERACHE, J., PEREIRA, G., IRIBARREN, J., SATTOLO, I. 2012. "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.

IERACHE. J, PEREIRA., G, NERVO., F, IRIBARREN, J.. 2013. Estado emocional centrado en estímulos, aplicando Interfase cerebro-máquina. XX CACIC, ISBN 978-987-3806-05-6.

JOYCE, R.; GUPTA, G. 1990. Identity authentication based on keystroke latencies. Commun. ACM 33, 2 (February 1990), 168-176. http://doi.acm.org/10.1145/75577.75582

GONZÁLEZ, N., CALOT , E.
Of Keystroke Dynamics In Free Text. En Biometrics Special Interest Group (Biosig), 2015 International Conference Of The, Páginas 1–5, Sep. 2015. ISBN 978-3-88579-639-8. DOI: 10.1109/Biosig.2015.7314606

GONZÁLEZ, N., CALOT, E. Y IERACHE, J... A Replication Of Two Free Text Keystroke Dynamics Experiments Under Harsher Conditions. En 2016 International Conference Of The Biometrics Special Interest Group (Biosig), Páginas 1–6, Sep. 2016. DOIi: 10.1109/Biosig.2016.7736905

LANG, P. J., BRADLEY, M. M. & CUTHBERT, B. N. INTERNATIONAL AFFECTIVE PICTURE SYSTEM (IAPS): TECHNICAL MANUAL AND AFFECTIVE RATINGS.NIMH CENT. STUDY EMOT. ATTEN. 39–58, HTTPS://DOI.ORG/10.1027/0269-8803/A000147 (1997)

LSIA: http://lsia.fi.uba.ar vigente marzo 2019

RODRÍGUEZ, J.M.; CALOT, E.P.; MERLINO H. 2014. "Clasificación de Prescripciones Médicas en Español". En "XX Congreso Argentino de Ciencias de la Computación". Nº 1258. ISBN 978-987-3806-05-6.

RODRÍGUEZ, J., MERLINO, H., GARCÍA-MARTÍNEZ, R.. 2015. Revisión sistemática comparativa de evolución de métodos de extracción de conocimiento para la web. Argentina. La Plata. Libro. Artículo Completo. CACIC 2015.

RODRÍGUEZ, J., MERLINO, H., PESADO, P., GARCÍA -MARTÍNEZ, R. 2016a. Performance Evaluation of Knowledge Extraction Methods. Suiza. Basel. 2016. Libro. Artículo Breve. Congreso. International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems (IEA-AIE 2016).

RODRÍGUEZ, J., MERLINO, H., PESADO, P., GARCÍA -MARTÍNEZ, R. 2016b. Clasificación de distintos conjuntos de datos utilizados en evaluación de métodos de extracción de conocimiento creados para la web. Argentina. San Luis. Congreso. CACIC XXII

RODRÍGUEZ, J., MERLINO, H., PESADO, P., GARCÍA -MARTÍNEZ, R. 2017a. Evaluation of Open Information Extraction Methods Using Reuters-21578 Database. Estados Unidos de América. New York. 2017. Libro. Artículo Completo. Congreso. 2nd International Conference on Machine Learning and Soft Computing (ICMLSC '18).

RODRÍGUEZ, J., MERLINO, H., GARCÍA -MARTÍNEZ, R.. 2017b. Automatic Characteristics Extraction for Sentiment Analysis Tasks. Suiza. Basel. 2017. Libro. Artículo Completo. Congreso. XXIII Congreso Argentino de Ciencias de la Computación (CACIC 2017). Universidad Nacional de La Plata.

RUSSELL, J. A circumplex model of affect. Journal of Personality and Social Psychology, 39(6):1161–1178, 1980. ISSN 0022-3514. DOI: 10.1037/h0077714

RUSSELL, J. Measures of emotion, páginas 83–111. Academic Press, 1989. ISBN 978-0-12-558704-4.

Métodos y herramientas para el análisis forense de dispositivos móviles

Susana Herrera, Liliana Figueroa, Daniel Ghunter, Cecilia Lara, Graciela Viaña, Analía Mendez, Norma Lesca

Instituto de Investigación en Informática y Sistemas de Información, Facultad de Ciencias Exactas y Tecnologías, Universidad Nacional de Santiago del Estero
sherrera@unse.edu.ar; lmvfigueroa@yahoo.com.ar; dgunther@unse.edu.ar; laraceciliacristina@gmail.com; gv857@hotmail.com; anmendez725@yahoo.com; norma.lesca@gmail.com

RESUMEN

Ante los desafíos que genera la evidencia digital en el sistema procesal penal como prueba en la investigación de delitos, la justicia necesita una regulación adecuada que permita obtener evidencias digitales legalmente aceptables, que ayuden a resolver conflictos según métodos científicos de recolección, análisis y validación. Aunque en algunas provincias de nuestro país existen guías de buenas prácticas y protocolos de actuación que orientan el trabajo pericial informático, Santiago del Estero no cuenta con normativa al respecto.

Como resultado del proyecto "Computación Móvil: desarrollo de aplicaciones y análisis forense", se propuso un conjunto de lineamientos a los que recurrir al momento de obtención de evidencia digital de dispositivos móviles, pero éste aún no se validó para su utilización como protocolo aprobado por el Ministerio Publico Fiscal y el Poder Judicial de la provincia.

Este proyecto pretende realizar investigación aplicada para validar el protocolo propuesto, así como el cumplimiento de estándares que garanticen la calidad de los procesos y sus resultados, tomando como referencia la familia de normas ISO/IEC 27000. Además, atendiendo requerimientos identificados durante el desarrollo del proyecto mencionado anteriormente, se investigará sobre herramientas tecnológicas de apoyo a la adquisición y gestión de evidencias digitales obtenidas de dispositivos móviles.

Palabras clave: Computación móvil, análisis forense, protocolo para el tratamiento de evidencias digitales.

CONTEXTO

En este artículo se presenta una propuesta de investigación que constituye una continuación del proyecto "Computación Móvil: desarrollo de aplicaciones y análisis forense", llevado adelante durante 2017 y 2018, financiado por el Consejo de Ciencia y Técnica de la Universidad Nacional de Santiago del Estero [4].

En dicho proyecto se lograron resultados referidos al análisis de la obtención legal de la evidencia digital en los códigos procesales de nuestro país, de antecedentes jurisprudenciales sobre tratamiento de evidencia digital en dispositivos móviles, investigación y análisis de protocolos vigentes en otras jurisdicciones, una propuesta de lineamientos para la obtención de evidencia digital en móviles en el ámbito del Ministerio Público Fiscal de Santiago del Estero y el estudio de repositorios que permitan la construcción de un modelo de datos para la gestión de las evidencias [15].

El mencionado protocolo, organizado en fases, abarca el proceso completo del tratamiento de la evidencia digital. Pone énfasis en las actividades y técnicas relacionadas con dispositivos móviles, constituyendo una herramienta para la planificación y control de dicho proceso en la investigación penal preparatoria. No obstante, se requiere de una revisión y evaluación que permita garantizar la validez

y eficacia probatoria de las evidencias digitales obtenidas en el proceso.

Por otro lado, se determinó que durante el examen de dispositivos móviles, los peritos informáticos se enfrentan a las siguientes dificultades:

1. No existe un software de informática forense que soporte la extracción de datos desde todos los dispositivos móviles existentes en el mercado.
2. Existe un gran número de programas diseñados para el sistema operativo Android donde los datos pueden ser interesantes para los investigadores de pruebas de penetración, pero hasta el momento no se cuenta con programas de apoyo de análisis forense de sus registros y datos.
3. Los requerimientos para los peritos forenses no están claros y bien definidos desde el inicio de la investigación.
4. Los delincuentes suelen eliminar los archivos de la memoria de sus dispositivos móviles, tratando de ocultar información sobre el crimen cometido.
5. Los laboratorios de informática forense no siempre pueden permitirse la compra de software especializado, debido a su alto costo.

En este contexto, se advierte la necesidad de trabajar en la validación del protocolo de actuación para el análisis de evidencias forenses obtenidas de dispositivos móviles propuesto anteriormente, y de incluir guías de actuación buscando aportar soluciones a los problemas que enfrentan los peritos informáticos.

1. INTRODUCCIÓN

Tradicionalmente, el concepto de evidencia se asoció al de evidencia física. Sin embargo, con el auge de las tecnologías de la información y la comunicación, surgió la evidencia digital, que también puede ser usada como medio de prueba en juicio [9].

En comparación con la evidencia tradicional, la digital es única: es sensiblemente frágil, anónima y volátil. Si fue presentada correctamente y su cadena de custodia no fue alterada, puede llegar a ser decisiva para resolver cualquier clase de delito.

La Informática Forense es una disciplina de las ciencias forenses que involucra la aplicación de metodologías científicas para identificar, preservar, recuperar, extraer, documentar e interpretar [16] evidencias procedentes de fuentes digitales con el fin de facilitar la reconstrucción de los hechos [11], para usar luego dichas evidencias como elemento probatorio en un proceso judicial [3,2].

Por otra parte, el empleo de dispositivos móviles se incrementó notablemente en los últimos diez años, debido a su facilidad de uso y a la propiedad de mantener en contacto permanentemente a sus usuarios. Esto generó un cambio significativo en el modo de comunicación de las personas, pero también incrementó su uso en actividades delictivas [8].

El análisis forense sobre dispositivos móviles es un campo relativamente nuevo, por lo que sus procedimientos y normas aún se encuentran en desarrollo [14]. Para lograr un análisis forense digital confiable y que la evidencia que se recoja sea admisible, auténtica, completa, fiable, entendible y creíble, es importante seguir los lineamientos formales establecidos en un protocolo de actuación.

Actualmente, la norma de alcance global de referencia para el análisis forense es la ISO/IEC 27037:2012 [6], aplicable a las organizaciones que necesiten establecer un marco formal de aceptabilidad. Dicha norma proporciona pautas para el manejo de la evidencia digital, sistematizando su identificación, recolección, análisis y preservación, y asegurando que los responsables de gestionar evidencia digital lo hagan con prácticas mundialmente aceptadas, de manera sistemática e imparcial. De acuerdo a ella, la evidencia digital es gobernada por los principios

fundamentales de relevancia, confiabilidad y suficiencia.

Por otro lado, para la adquisición de datos y clasificación de herramientas, es posible tomar como referencia el "Sistema de clasificación de herramientas forenses de dispositivos móviles" [1], que presenta la Pirámide Móvil Forense con cinco niveles de extracción y análisis requeridos en cada dispositivo móvil, teniendo en cuenta la solicitud y los detalles de la investigación.

En relación a la problemática referida, las principales contribuciones desarrolladas en el ámbito de nuestro país son:

- *"Guía de obtención, preservación y tratamiento de la evidencia digital"* [13] de la Unidad Fiscal Especializada en Ciberdelitos de la Procuración General de la Nación. Este trabajo brinda una serie recomendaciones utilizadas a nivel mundial y de herramientas de investigación, abordando el modo en el que se debe obtener, observar y tratar la evidencia digital para mejorar los niveles de eficiencia en persecución penal.
- *"Protocolo de actuación para pericias informáticas del Poder Judicial de Neuquén"* [10], donde se aborda la pericia sobre telefonía celular como parte de la especialidad de la informática forense, un procedimiento para pericias informáticas sobre telefonía celular y el uso del Universal Forensic Extraction Device (UFED).
- *"Guía integral de empleo de la informática forense en el proceso penal"*[5], del Laboratorio de Investigación y Desarrollo de Tecnología en Informática Forense (InfoLab), que presenta los aspectos básicos a considerar en las labores de búsqueda, obtención, preservación, examen pericial y presentación de evidencias digitales en el proceso penal, a fin de garantizar su validez y eficacia probatoria.
- *"Guía de procedimientos para pericias de dispositivos móviles"* [12] del Poder Judicial de Río Negro, donde se presenta la metodología del ciclo de vida de la evidencia digital, considerando en estado en que se encuentra el equipo y el lugar donde se lleva a cabo la intervención.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Considerando la amplitud de los aspectos relacionados con la informática forense sobre computación móvil, el presente proyecto de investigación se refiere a:

Informática Forense: métodos y herramientas para el análisis forense de dispositivos móviles.

A partir de él, 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.

- Repositorio digital para la gestión de evidencias digitales extraídas de dispositivos móviles.

3. OBJETIVOS

El objetivo general de la investigación propuesta es:

- *Contribuir al mejoramiento de la calidad del proceso de obtención de evidencias digitales obtenidas de dispositivos móviles en el ámbito judicial de Santiago del Estero.*

Los objetivos específicos que permitirán alcanzar el objetivo general son:

- *Validar el Protocolo para la obtención de evidencias digitales de dispositivos móviles en el ámbito judicial de Santiago del Estero.*

- *Ampliar el Protocolo de obtención de evidencias digitales de dispositivos móviles considerando el sistema de clasificación de*

herramientas forenses de dispositivos móviles.

- *Analizar alternativas de construcción de repositorio digital para la gestión de evidencias digitales extraídas de dispositivos móviles.*

En esta propuesta se pretende llevar a cabo una investigación aplicada para validar el Protocolo de Actuación propuesto en el proyecto "Computación Móvil: desarrollo de aplicaciones y análisis forense", para el cumplimiento de buenas prácticas tendientes a garantizar la calidad de los procesos aplicados y los resultados obtenidos, tomando como referencia la familia de normas ISO/IEC 27000, más específicamente las normas ISO/IEC 27037:2012 [6] y 27042:2015 [7].

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, la variable a estudiar es la "optimización de la gestión de evidencias criminales obtenidas de dispositivos móviles", que podrá ser evaluada a través de indicadores cuantitativos que se pueden aplicar a casos de prueba especialmente diseñados.

Durante la validación de la aplicabilidad del protocolo propuesto se realizarán experiencias de puesta en marcha con el propósito de lograr un consenso técnico en la materia, y permitiendo que se generen nuevas versiones del mismo a partir de las sugerencias y recomendaciones derivadas de los informes de estas experiencias.

Se considera que la validación del mencionado protocolo, y su posterior aceptación por parte del Ministerio Público Fiscal y el Poder Judicial de Santiago del Estero, traerán un beneficio muy importante para la justicia santiagueña, dado que actualmente no existe un procedimiento claro y definido al respecto. Esto permitiría mejorar la calidad de las evidencias digitales y ayudará en la labor de los fiscales y peritos de la provincia.

4. FORMACIÓN DE RECURSOS HUMANOS

La Directora y Codirectora del proyecto pertenecen al Departamento de Informática de la Universidad Nacional de Santiago del Estero. El asesor es experto en Informática Forense y jefe del área de Informática Forense del Poder Judicial de Río Negro.

Los investigadores constituyen un equipo interdisciplinario conformado por cuatro docentes de la UNSE y un investigador externo, con profesión en Informática, Electromecánica y Derecho. Estos poseen distintas categorías de investigación y algunos desempeñan sus actividades profesionales en el Poder Judicial de la Provincia de Santiago del Estero y el Gabinete de Ciencias Forenses del Ministerio Público Fiscal de Santiago del Estero.

El equipo de investigación se encuentra asistiendo y asesorando a alumnos de grado y posgrado de UNSE que realizan sus trabajos de finalización de carrera en temáticas relacionadas con esta línea de investigación, los que se encuentran en la etapa de propuesta inicial.

5. REFERENCIAS

1. AYERS, R.; BROTHERS, S.; JANSEN, W. (2014). Guidelines on Mobile Device Forensics. NIST Special Publication 800-101. Revision 1. http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-101r1.pdf
2. CASTILLO, C.; ROMERO, A.; CANO, J. (2008). Análisis Forense Orientado a Incidentes en Teléfonos Celulares GSM: Una Guía Metodológica. XXXIV Conferencia Latinoamericana de Informática, Centro Latinoamericano de Estudios en Informática (CLEI). http://www.clei2008.org.ar/

3. DEL PINO, S. (2007). Introducción a la informática forense. Pontificia Universidad Católica del Ecuador. http://www.criptored.upm.es/guiateoria/gt_m592b.htm

4. FENNEMA, M.; FIGUEORA, L.; VIAÑA, G.; LESCA GOMEZ, N.; LARA, C. (2017). Tratamiento de evidencias digitales forenses en dispositivos móviles. XIX Workshop de Investigadores en Ciencias de la Computación. ISBN 978-987-42-5143-5.

5. INFO-LAB. (2015). Guía integral de empleo de la informática forense en el proceso penal. http://info-lab.org.ar/images/pdf/14.pdf

6. 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

7. ISO/IEC 27042:2015 (en) Information technology — Security techniques — Guidelines for the analysis and interpretation of digital evidence. https://www.iso.org/obp/ui/#iso:std:iso-iec:27042:ed-1:v1:en

8. MELLAR, B. (2004). Forensic Examination of Mobile Phones. Digital Investigation – Elsevier. United Kingdom. http://faculty.colostate-pueblo.edu/dawn.spencer/Cis462/Homework/Ch4/Forensic%20examination%20of%20mobile%20phones.pdf

9. ORTA MARTINEZ, R. (2013). Informática Forense como Medio de Pruebas. http://www.dragonjar.org//informatica-forense-como-medio-de-prueba.xhtml

10. PODER JUDICIAL DE NEUQUÉN. (2013). Pericias informáticas sobre telefonía celular. http://200.70.33.130/images2/Biblioteca/ProtocoloPericiasTelefoniaCelular.pdf

11. REITH, M.; CLINT, C.; GUNSCH G. (2002). An Examination of Digital Forensic Models. International Journal of Digital Evidence, Air Force Institute of Technology, Volume 1 Issue 3. www.utica.edu/academic/institutes/ecii/publications/articles/A04A40DC-A6F6-F2C1-98F94F16AF57232D.pdf.

12. SUP. TRIBUNAL DE JUSTICIA DEL PODER JUDICIAL DE RÍO NEGRO. (2015). Guía de procedimientos para pericias de dispositivos móviles. http://digesto.jusrionegro.gov.ar/bitstream/handle/123456789/455/Ac011-15.pdf?sequence=1&isAllowed=y

13. UNIDAD FISCAL ESPECIALIZADA EN CIBERDELITOS. (2016). Guía de obtención, preservación y tratamiento de evidencia digital. http://www.fiscales.gob.ar/procuracion-general/wp-content/uploads/sites/9/2016/04/PGN-0756-2016-001.pdf

14. VARSALONE, J., KUBASIAK, R. (2009). Mac Os X, iPod and iPhone Forensic Analysis DVD Toolkit. Syngress Publishing, Inc, pp. 355-475.

15. VIAÑA, G.; FIGUEORA, L.; LARA, C.; LESCA GOMEZ, N. (2018). Protocolo de actuación para recolección y preservación de la evidencia digital móvil en el Sistema Procesal Penal de Santiago del Estero. VI Congreso Nacional de Ingeniería en Informática y Sistemas de Información. ISSN 2347-0372.

16. ZDZIARSKI, J. (2008). iPhone Forensics, Recovering Evidence, Personal Data & Corporate Assets. O"Reilly Media, Inc.

Orquestación de recursos para realizar cómputo de big data sobre arquitecturas distribuidas

María Murazzo[1], Nelson Rodríguez[1], Miguel Guevara[1], Martin Tello[4], Matías Alonso[5], Diego Medel[1], Jorge Mercado[2], Pablo Gómez[5], Fabiana Picolli[3]

[1] Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.
[2] Departamento de Matemática, Facultad de Ingeniería - U.N.S.J.
[3] Departamento de Informática - F.C.F.M. y N, Universidad Nacional de San Luis
[4]Alumno Avanzado Licenciatura en Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.
[5] Alumno Avanzado Licenciatura en Sist. de Información - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Ignacio de la Roza y Meglioli. C.P. 5402. Rivadavia. San Juan, 0264 4234129

marite@unsj-cuim.edu.ar, nelson@iinfo.unsj.edu.ar, migueljoseguevaratencio@gmail.com, martinl.tello@gmail.com, matialonso95@gmail.com, vdiego.unsj@hotmail.com, jnmercado@unsj-cuim.edu.ar, mpiccoli@unsl.edu.ar

Resumen

Los avances tecnológicos, en particular IoT, han favorecido la generación de grandes cantidades de datos, los cuales necesitan ser almacenados y procesados de manera eficiente. Surge así el paradigma Big Data, donde el principal requerimiento no solo es la capacidad de cómputo, sino el manejo en un tiempo razonable de ingentes cantidades de datos. En este contexto, las aplicaciones para big data necesitan ser escalables, livianas, autocontenidas, distribuidas y replicadas con el objetivo de lograr la mejor performance frente a variaciones del volumen de datos.

Para lograr esto, esta línea de investigación propone ajustar la construcción de aplicaciones a la filosofía DevOps y una arquitectura basada en microservicios los cuales puedan ser implementados con contenedores. La replicación y distribución para lograr altos niveles de escalabilidad se plantea mediante la orquestación de contenedores sobre una arquitectura distribuida, la cual puede ser física o virtual. De esta manera será posible desarrollar aplicaciones que escalen de forma elástica en función de los requerimientos de la aplicación misma, independiente del software y hardware subyacente.

Palabras clave: *Big Data, Micrsoservices, Conteiners, DevOps, Distributed Systems.*

Contexto

El presente trabajo se encuadra dentro del área de I/D Innovación en Sistemas de Software y se enmarca dentro del proyecto de investigación: Orquestación de servicios para la Continuidad de Edge al Cloud, que ha sido aprobado y está en desarrollo para el período 2018-2019. Asimismo el grupo de investigación viene trabajando en proyectos relacionados con la computación móvil, distribuida y de alta performance desde hace más de 18 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 integración de Cloud Computing (CC) [1] con Internet of Thing (IoT) [2] representa el próximo gran paso en las TI del futuro. Las aplicaciones que surjan de esta integración abrirán nuevas perspectivas de negocio y oportunidades de investigación. Gracias al paradigma CloudIoT [3], la vida cotidiana mejorará, por ejemplo, en las Smart city [4], permitirán servicios públicos más eficientes través de aplicaciones ubicuas que mejorarán la calidad de vida de los ciudadanos. El futuro

de este paradigma pasa por la convergencia hacia una plataforma de servicio común que supere los desafíos planteados por la heterogeneidad de los dispositivos y las tecnologías involucradas. Esto conlleva tener en cuenta la ubicuidad y la omnipresencia de los dispositivos, lo cual, requieren plataformas de computación escalables capaces de asegurar la continuidad hacia el cloud [5] .

La aceleración en la tasa de generación de datos a dando lugar al paradigma Big Data [6], el cual se caracteriza por tres aspectos: a) el volumen de los datos, b) los datos no se pueden almacenar de forma tradicional y c) los datos son generados, capturados y procesados con rapidez. Estas características hacen que los sistemas de cómputo, las metodologías de desarrollo y los framework de desarrollo convencionales sean inapropiados para lograr una adecuada gestión del big data [7].

Construir aplicaciones para big data tiene como principal problema la necesidad de administrar recursos en forma elástica, lo cual implica analizar y planificar la forma en la cual la aplicación debe escalar. Tradicionalmente, la construcción de aplicaciones ha seguido un paradigma de escalado vertical mediante el cual la solución a los problemas de cuellos de botellas se resuelve mediante el aprovisionamiento de mayor cantidad de recursos. Pero cuando se debe almacenar, procesar y analizar grandes cantidades de datos, las aplicaciones deben pensarse y construirse de forma distribuida; obligando a contar con mecanismos de escalabilidad horizontal, que permita distribuir la carga de trabajo dinámicamente y paralelizar las tareas [8] [9] .

La escalabilidad horizontal implica balanceo de carga, replicación de recursos, reinstanciación de recursos en tiempo real y optimización en su uso, de forma que el usuario no perciba degradación de performance [10]. Para lograr esto es necesario migrar la forma en la cual se piensan y se construyen las aplicaciones con el objetivo que se alineen con la misión y la visión organizacional.

Arquitectura de Software

Tradicionalmente, el desarrollo de software se ha ajustado en una arquitectura monolítica, donde las aplicaciones poseen una única base de código, la cual ofrece una variedad de servicios, que utilizan diferentes interfaces tales como páginas HTML encapsuladas a un mismo repositorio. La desventaja que presenta este enfoque es que, si se necesita realizar una modificación sobre una funcionalidad, impactará en las demás funcionalidades del sistema de modo tal que aumenta el costo de llevar a cabo dicha tarea. Otra desventaja es que, al tratarse de un esquema único, pueden generar un cuello de botella debido a que un error en la aplicación deja sin servicios todo el sistema.

Trabajar con grandes volúmenes de datos provenientes del IoT con las restricciones generadas por la necesidad de asegurar la continuidad al cloud, hacen inviable aplicar este modelo arquitectónico. Es por ello que resulta apropiado trabajar con arquitecturas basadas en servicios. En este caso, cada servicio son aplicaciones de negocio auto descriptivas y modulares que exponen la lógica organizacional como servicios a través de Internet por medio de una interfaz programable y donde el protocolo de Internet (IP) puede ser utilizado para encontrar e invocar esos servicios. Un servicio es el elemento clave en la integración de diferentes sistemas de información, ya que los sistemas de información pueden basarse en diferentes plataformas, lenguajes de programación y tecnologías. En particular, la arquitectura de microservicios [11] permite construir aplicaciones distribuidas como un conjunto de pequeños servicios desacoplados, desplegados de manera independiente y que trabajen coordinadamente.

Una de las características más importantes que provee los microservicios es la escalabilidad [12]. Debido a que los microservicios se desarrollan en forma independiente uno de otros, también escalan en forma independiente, esto es muy importante cuando se trabaja con grandes volúmenes de datos pues permite la implementación de carga compartida de trabajo.

Los microservicios dependen no solo de la tecnología que se está configurando, sino también de una organización que tenga la cultura, el conocimiento y las estructuras establecidas para que los equipos de desarrollo puedan adoptar este modelo. Los microservicios forman parte de un cambio más amplio en los departamentos de TI hacia una cultura DevOps [13], en la que los equipos de desarrollo y operaciones trabajan estrechamente para respaldar una aplicación a lo largo de su ciclo de vida y pasan por un ciclo de publicación rápido o incluso continuo en lugar de un ciclo tradicional largo.

Contenerización

El uso de microservicios implica la construcción de aplicaciones a partir de múltiples componentes autosuficientes. Estos componentes pueden ser implementados con contenedores [14]. Un contenedor, es un paquete de software que, internamente, contiene la aplicación que se quiere ejecutar y todas sus dependencias usando indirectamente el kernel del sistema operativo subyacente. Los contenedores permiten implementar aplicaciones distribuidas debido a dos características: portabilidad, pues los contenedores se pueden ejecutar de forma independiente al hardware y al software donde se crearon; y baja sobrecarga, pues son livianos e introducen menos overhead que las VM (Virtual Machine) [15].

La principal ventaja [16] de utilizar contenedores es que no depende de las instalaciones del sistema anfitrión, es decir es posible realizar las instalaciones necesarias para desplegar aplicaciones dentro de un contenedor. Otra ventaja es que permite la portabilidad en las aplicaciones, esto quiere decir que además de crear contenedores, es posible definir repositorios que eventualmente alojarán a los contenedores con el objeto que otros usuarios los descarguen y los consuman como servicio.

Quizás la características más relevante de los contenedores es que permiten desplegar aplicaciones de forma rápida y escalables, ya que permite replicar ambientes de desarrollo, prueba y producción en poco tiempo, pudiendo aumentar o disminuir la cantidad de contenedores a medida que sean necesitados (escalabilidad elástica) [17].

El uso conjunto de contenedores y microservicios [18] mejora las capacidades de escalado, ya que el microservicio es portable y reutilizable; mientras que los contenedores proporcionan recursos eficientes, principalmente el empaquetado de aplicaciones y sus dependencias en un contenedor virtual que puede ejecutarse en cualquier servidor.

Orquestación de Servicios

Se ha dicho que cuando se trabaja con datos masivos es necesario escalar horizontalmente las aplicaciones, esto se puede lograr mediante el uso de arquitecturas distribuidas con el objeto de repartir los datos y la carga de trabajo. Para lograr esta forma de distribución es fundamental abstraerse de la complejidad de la plataforma subyacente, lo cual se logra mediante el uso de una arquitectura de hardware distribuida donde implementar contenedores que interactúen entre si [19].

En esta forma de trabajar se necesita administrar y coordinar los contenedores (microservicios), para poder lograrlo se realiza la orquestación o coreografía de los mismos [20].

Orquestar significa que hay una entidad central (idealmente un microservicio en sí mismo) que coordina la comunicación entre los microservicios. Mientras que para microservicios coreografiados, cada servicio implicado en dicha coreografía "conoce" exactamente cuándo ejecutar sus operaciones y con quién debe interactuar, lo cual le quita transparencia e independencia a los servicios. Los servicios orquestados no "conocen" (y no necesitan conocer) que están implicados en un proceso de composición y que forman parte de un proceso de negocio de nivel más alto. Solamente el coordinador central de la orquestación es "consciente" de la meta a conseguir, por lo que la orquestación se centraliza mediante definiciones explícitas de las operaciones y del orden en el que se deben invocar los servicios.

Gracias a la orquestación de contenedores [21] se puede trabajar de forma transparente frente al aumento de carga de trabajo, falla de algún nodo, creación de nuevos contenedores, etc. El orquestador decide cuándo adaptar el sistema (por ejemplo, para iniciar la ejecución de un contenedor) y donde la adaptación tendrá lugar (por ejemplo, en qué nodo se ejecutarán los contenedores). Orquestar contenedores [22], aísla al usuario de los detalles de implementación, lo cual provee transparencia en el acceso y ejecución de los contenedores. Este aspecto es importante cuando se desea implementar replicación y distribución de datos masivos con el objetivo de lograr balanceo de carga y auto escalado horizontal, dependiente de la sobrecarga del volumen de información. Además, la orquestación, permitirá un aprovisionamiento de recursos bajo demanda y un escalado horizontal de la aplicación desarrollada.

Desarrollar aplicaciones de esta manera permitirá mitigar los problemas de continuidad de los datos provenientes del IoT para logar un procesamiento eficiente. Además, gracias a la portabilidad será posible alojar las aplicaciones en cualquier lugar desde la fuente de generación de los datos hasta el extremo (edge).

Líneas de Investigación, Desarrollo e Innovación

Las tareas de investigación se centrarán en la generación de buenas prácticas tendientes a lograr desarrollo de aplicaciones para big data, provenientes del IoT. Esto incluye el análisis de:

- Mejores prácticas de desarrollo basada en la filosofía DevOps.
- Métodos y herramientas de pre procesamiento y almacenamiento de grandes volúmenes de datos.
- Herramientas de contenerización de datos masivos.
- Definición de middleware de orquestación sobre arquitecturas distribuidas.
- Herramientas de medición de escalabilidad.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos dieciocho años se trabajó en una variedad de sistemas distribuidos y paralelos. Ya sea en 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, como así también con arquitecturas de menores prestaciones (computación móvil y Edge Computing).

Dicha experiencia motivó la elaboración del proyecto de investigación del que forma parte la presente propuesta. El grupo ha realizado publicaciones en el área durante los últimos años: dieciocho trabajos de investigación en diferentes Congresos y Jornadas, se realizaron tres publicaciones en revistas científicas y se transfirieron los resultados mediante seis conferencias en eventos científicos y encuentros de divulgación.

Se han aprobado diecisiete tesinas de grado y un trabajo de especialización y se incorporaron dos becarios de investigación categoría alumno, un becario doctoral CONICET y se encuentran en desarrollo tesinas de grado y de maestría.

Objetivos

El objetivo del grupo de investigación es definir y construir una capa de orquestación de recursos, alineada con la filosofía DevOps, para brindar servicios de almacenamiento y procesamiento de grandes volúmenes de datos sobre arquitecturas distribuidas usando la tecnología de contenedores. Esta capa de orquestación deberá asegurar la continuidad mediante técnicas de replicación, distribución y balanceo de carga con el fin de maximizar la eficiencia de las aplicaciones.

Formación de Recursos Humanos

El equipo de trabajo está compuesto por los docentes-investigadores de la línea de investigación presentada que figuran en este trabajo. Cabe destacar que el proyecto marco de la presente línea de investigación incluye

investigadores de UNSL, Champagnat (Mendoza) y UNLaR.

Con respecto a la formación de recursos humanos, se está realizando una tesis doctoral, una tesis de maestría y cuatro tesinas de grado. Además este año se cuenta con un alumno el cual ha accedido a la beca CIN y cuyo tema se enmarca en la presente línea de investigación.

Por otro lado, se espera aumentar el número de publicaciones y también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o actividades de divulgación y asesoramiento a empresas y organismos del estado.

Referencias

[1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," 2011.

[2] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, "Internet of Things (IoT): A vision, architectural elements, and future directions," *Futur. Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.

[3] S. M. Babu, A. J. Lakshmi, and B. T. Rao, "A study on cloud based Internet of Things: CloudIoT," in *2015 Global Conference on Communication Technologies (GCCT)*, 2015, pp. 60–65.

[4] A. Cocchia, "Smart and Digital City: A Systematic Literature Review," Springer, Cham, 2014, pp. 13–43.

[5] N. R. Rodríguez *et al.*, "Orquestación de servicios para la continuidad edge al cloud," 2018.

[6] A. Katal, M. Wazid, and R. H. Goudar, "Big data: Issues, challenges, tools and Good practices," in *2013 Sixth International Conference on Contemporary Computing (IC3)*, 2013, pp. 404–409.

[7] M. Parashar, X. Li, and Wiley InterScience (Online service), *Advanced computational infrastructures for parallel and distributed adaptive applications*. John Wiley & Sons, 2010.

[8] M. A. Murazzo, M. J. Guevara, M. Tello, N. R. Rodríguez, F. Piccoli, and M. Giménez, "Orquestación de servicios para el desarrollo de aplicaciones para big data," 2018.

[9] M. Barrionuevo *et al.*, "Estrategias y análisis orientados al manejo de datos masivos usando computación de alto desempeño," 2018.

[10] I. Gorton and J. Klein, "Distribution, Data, Deployment: Software Architecture Convergence in Big Data Systems," *IEEE Softw.*, vol. 32, no. 3, pp. 78–85, May 2015.

[11] K. Bakshi, "Microservices-based software architecture and approaches," in *2017 IEEE Aerospace Conference*, 2017, pp. 1–8.

[12] N. Dragoni, I. Lanese, S. T. Larsen, M. Mazzara, R. Mustafin, and L. Safina, "Microservices: How To Make Your Application Scale," Springer, Cham, 2018, pp. 95–104.

[13] L. Zhu, L. Bass, and G. Champlin-Scharff, "DevOps and Its Practices," *IEEE Softw.*, vol. 33, no. 3, pp. 32–34, May 2016.

[14] N. Kratzke, "About Microservices, Containers and their Underestimated Impact on Network Performance," Sep. 2017.

[15] R. Morabito, J. Kjallman, and M. Komu, "Hypervisors vs. Lightweight Virtualization: A Performance Comparison," in *2015 IEEE International Conference on Cloud Engineering*, 2015, pp. 386–393.

[16] C. Ruiz, E. Jeanvoine, and L. Nussbaum, "Performance Evaluation of Containers for HPC," Springer, Cham, 2015, pp. 813–824.

[17] S. R. Brus, D. Wirasaet, J. J. Westerink, and C. Dawson, "Performance and Scalability Improvements for Discontinuous Galerkin Solutions to Conservation Laws on Unstructured Grids," *J. Sci. Comput.*, vol. 70, no. 1, pp. 210–242, Jan. 2017.

[18] S. Yangui, M. Mohamed, S. Tata, and S. Moalla, "Scalable Service Containers," in *2011 IEEE Third International Conference on Cloud Computing Technology and Science*, 2011, pp. 348–356.

[19] C. Pahl and B. Lee, "Containers and Clusters for Edge Cloud Architectures -- A Technology Review," in *2015 3rd International Conference on Future Internet of Things and Cloud*, 2015, pp. 379–386.

[20] N. Busi, R. Gorrieri, C. Guidi, R. Lucchi, and G. Zavattaro, "Choreography and Orchestration: A Synergic Approach for System Design," Springer, Berlin, Heidelberg, 2005, pp. 228–240.

[21] W. Gerlach, W. Tang, A. Wilke, D. Olson, and F. Meyer, "Container Orchestration for Scientific Workflows," in *2015 IEEE International Conference on Cloud Engineering*, 2015, pp. 377–378.

[22] J. Rufino, M. Alam, J. Ferreira, A. Rehman, and K. F. Tsang, "Orchestration of containerized microservices for IIoT using Docker," in *2017 IEEE International Conference on Industrial Technology (ICIT)*, 2017, pp. 1532–1536.

Prácticas ingenieriles aplicadas para la implementación de sistemas inteligentes basados en machine learning

Pollo-Cattaneo, Ma. F; Pytel, P.; Vegega, C.; Ramón, H.; Deroche, A.; Straccia, L.; Velazquez, G., Orozco-Gonzalez, M. & Panizza, L.

Grupo de Estudio en Metodologías de Ingeniería de Software. Facultad Regional Buenos Aires. Universidad Tecnológica Nacional. Argentina.
{flo.pollo, ppytel}@gmail.com

Resumen

La Inteligencia Artificial (IA) tiene el potencial de aportar en un futuro cercano grandes beneficios a la sociedad, gracias a la aplicación de Sistemas Inteligentes implementados mediante el uso de algoritmos de Machine Learning (o Aprendizaje Automático). No obstante, para aplicar Machine Learning, es imprescindible conocer varios aspectos relacionados al problema, la organización y su contexto, así como realizar un reconocimiento de las fuentes de información disponibles identificando su calidad. En caso contrario, el sistema podría estar resolviendo un problema diferente del que se quiere dar solución pudiendo provocar a largo plazo situaciones de sexismo, racismo y otras formas de discriminación En este contexto, el trabajo propuesto se inscribe en una línea de investigación que busca adaptar prácticas ingenieriles en la implementación de Sistemas Software Inteligentes basados en Machine Learning para evitar los problemas antes mencionados.

Palabras clave:
Sistema Inteligente, Inteligencia Artificial, Machine Learning, Prácticas Ingenieriles, Ingeniería de Software, Ingeniería en Sistemas de Información.

Contexto

El trabajo propuesto se desarrolla en el marco de las actividades del Grupo de Estudio de Metodologías para Ingeniería en Software y Sistemas de Información (GEMIS), el cual se encuentra conformado por un equipo de docentes y alumnos dentro del ámbito de la Facultad Regional Buenos Aires de la Universidad Tecnológica Nacional (UTN-FRBA). Este grupo busca la sistematización de cuerpos de conocimientos y promoción de sus aplicaciones y abordajes metodológicos en todo tipo de escenarios (convencionales y no convencionales). Tal es el caso de la implementación de un Sistema Inteligente para asistir a docentes y alumnos de una asignatura de grado en sus tareas tanto dentro como fuera del aula [1]. Sin embargo, a pesar de que se ha avanzado en varios módulos del sistema propuesto [2-5], el grado de avance ha sido más lento de lo esperado inicialmente debido a complicaciones asociadas a la definición de los conocimientos necesarios y la disponibilidad de datos representativos con calidad suficiente para entrenar a los algoritmos correspondientes. En este sentido, se han propuesto y aplicado métodos ad-hoc para solucionar este tipo de dificultades. Como es factible que muchos de los métodos puedan ser aplicados en otros ámbitos, se considera

de utilidad intentar unificarlos para facilitar el trabajo de los ingenieros.

Líneas de Investigación, Desarrollo e Innovación

El éxito de grandes empresas tecnológicas en la adopción de la Inteligencia Artificial (IA) ha llevado a que empresas de todo el mundo también adopten dichas tecnologías [6]. Esta situación conlleva a pronosticar que los ingresos derivados de la IA experimenten un incremento de u$s 37.000 millones del 2018 al 2025 [7]. Por otro lado, instituciones reconocidas, como la Universidad de Standford [8], publican reportes que destacan la importancia que tendrá la IA en un futuro cercano. Cabe destacar que esos beneficios están asociados al Aprendizaje Automático, más conocido como Machine Learning por su nombre en inglés. Esta disciplina busca estudiar y modelar los procesos de aprendizaje, con sus múltiples manifestaciones, para poder ser trasladados a las computadoras [9; 10] y facilitar, de esta manera, la implementación de sistemas Software en ámbitos complejos. Los sistemas Software así generados se designan Sistemas Inteligentes [11] y, pueden ser aplicados para resolver diferentes tipos de problemas [12]. Por lo tanto, dependiendo del problema que se desee resolver es posible utilizar alguno de los diversos algoritmos disponibles provistos por Machine Learning [11-14], por lo que una de las primeras dificultades asociadas a la construcción de los Sistemas Inteligentes tiene que ver con determinar cuál aplicar. De acuerdo al Teorema de 'No Free Lunch' [15], no existe ningún algoritmo que pueda ser aplicado a cualquier tipo de problema. Cada uno tiene sus puntos fuertes y débiles, que deberán ser contrastados contra los elementos del problema para determinar cuál es la mejor solución a aplicar. En otras palabras, siempre se debe seleccionar el algoritmo teniendo en cuenta los objetivos planteados, las características del dominio y los datos disponibles.

En este sentido, es interesante resaltar esta doble relación entre los datos y la tecnología a aplicar [11]. Para seleccionar la mejor tecnología, es útil conocer las características de los datos disponibles, pero para definir dichos datos se debe conocer la tecnología a aplicar. Esto se debe a que todos los algoritmos antes mencionados son métodos de caja negra y 'basados en datos' [9-10], por poder generar automáticamente modelos no-lineales sobre las relaciones entre los datos suministrados. De esta manera, se los pueden emplear para resolver un problema sin necesitar que se programe manualmente ninguna lógica.

Por consiguiente, otra de las cuestiones principales, asociadas al uso de Machine Learning, tiene que ver con obtener los datos necesarios para ser suministrados al algoritmo, lo cual no es una tarea trivial. Por ejemplo, tómese la cuestión de determinar la cantidad de información histórica que se necesita para producir los mejores resultados [15]. Según [16], la respuesta habitual a la pregunta "¿cuántos datos se necesitan?" es "lo más posible" dado que cuantos más datos se tenga, mejor se podrá identificar la estructura del modelo. No obstante, al llevar a cabo un proyecto real, es imprescindible ponerle algún límite a la cantidad de datos que se vayan a aplicar [12]. Asimismo, las propuestas existentes se consideran excesivamente simplificadas debido a que ignoran aspectos como: la variabilidad aleatoria subyacente de los mismos o, las características del dominio del problema. Entonces, es necesario primero identificar las fuentes de datos disponibles y

comprender sus características. Sólo así será posible recolectar datos suficientemente representativos del problema que se quiere resolver.

Cuando los datos no son suficientemente representativos se dice que los mismos presentan un Sesgo (o Bias en inglés). Si los datos presentan un sesgo, entonces se corre el riesgo de generar un Sistema Inteligente que no se basa en la realidad y pueda producir resultados erróneos [12]. En otras palabras, se podría estar entrenando a los algoritmos para resolver un problema diferente del que se quiere resolver. Entonces, es imprescindible conocer de antemano los sesgos asociados a los datos y al Sistema Inteligente de forma que puedan ser comprendidos por sus futuros usuarios para evitar malos entendidos y situaciones de discriminación [17]. No es raro que desarrolladores con las mejores intenciones puedan producir inadvertidamente Sistemas Inteligentes con resultados prejuiciosos, porque incluso ellos pueden no entender lo suficiente al problema, su contexto y los datos como para prevenir resultados no intencionados [18]. Lo peor de este escenario es que el sesgo puede ser tan sutil que no se detecte durante las pruebas. Si luego ese sistema se pone en operación y los usuarios llegan a confiar ciegamente en los resultados a largo plazo, se podrían provocar situaciones de sexismo, racismo y otras formas de discriminación [19].

Finalmente, además de las dificultades antes mencionadas, también existen otros factores que afectarán el rendimiento de un Sistema Inteligente, entre los que se destaca la configuración (o 'parametrización') del algoritmo a ser aplicado. Aunque existen diversos mecanismos que permiten definir automáticamente estas configuraciones, en ciertas implementaciones se considera de gran utilidad también incorporar modificaciones basadas en el conocimiento de expertos del dominio.

En este contexto, se desprende que las principales dificultades asociadas a la implementación de un Sistema Inteligente dependen de aspectos relativos al problema que se quiere resolver. En tal sentido, es imprescindible conocer las expectativas que el cliente tiene sobre el modelo que se traducirán en los objetivos que deberá cumplir. Dado que los mismos normalmente estarán muy relacionados con las metas estratégicas y tácticas de la organización, también se deben entender características propias de la organización, su contexto y los procesos de negocio realizados. Una vez que estos elementos son identificados, es necesario realizar un reconocimiento inicial de las fuentes de información disponibles en la organización identificando cuáles fuentes se encuentran informatizadas (en repositorios de datos) y cuáles no. En el caso de los repositorios, es muy importante también determinar la estructura, naturaleza y calidad de los datos disponibles. Por consiguiente, se considera indispensable elicitar todos estos requerimientos en forma completa y detallada para poder construir así el modelo que satisfaga las necesidades predictivas del cliente.

Dentro de la Ingeniería de Software existe la Ingeniería de Requerimientos, que es la encargada de recopilar, analizar y validar los deseos y necesidades de los distintos stakeholders sobre el sistema Software a implementar [20]. Sin embargo, a diferencia de los proyectos de desarrollo de Software tradicional, la problemática abordada en la implementación de Sistemas Inteligentes es diferente. Por lo tanto, no es posible aplicar sus métodos y técnicas directamente y deben ser adaptadas previamente.

A partir de un relevamiento efectuado en el campo metodológico, se han encontrado diversas propuestas vinculadas a la definición de estos elementos pero la mayoría están más orientadas a la identificación de aspectos relativos de una tecnología específica, dejando de lado las particularidades de otros algoritmos. Además, cada propuesta se concentra en algún aspecto dejando de lado otros por lo que no existe una guía integral para el desarrollo de esta arquitectura. En consecuencia, se considera de interés unificar las propuestas existentes.

Resultados y Objetivos

El proyecto propuesto se inscribe en una línea de investigación dentro del ámbito de la Ingeniería en Sistemas de Información (en general) y, la Ingeniería de Software (en particular), para buscar sistematizar el cuerpo de conocimiento existente y así sentar las bases del desarrollo de Prácticas Ingenieriles en la implementación de Sistemas Inteligentes basados en Machine Learning. Por lo cual, se considera tanto adaptar procesos, métodos, técnicas y herramientas existentes, así como también desarrollar nuevas, para ser utilizadas en la construcción de Sistemas Software Inteligentes. Para ello, se propondrá un Modelo de Proceso que englobe un conjunto de procesos, métodos, técnicas y herramientas que van a ser aplicados durante todas las fases del ciclo de vida de un proyecto de este tipo. De esta manera, se busca continuar y ampliar los trabajos desarrollados por GEMIS en proyectos anteriores.

Asociados a este objetivo general se definen los siguientes objetivos específicos:

- Identificar los métodos, técnicas y herramientas para ser utilizados en la elicitación de requerimientos del Sistema Inteligente.
- Determinar las tareas de descubrimiento, evaluación, recolección y preparación de los datos necesarios para la construcción del Sistema Inteligente.
- Definir un método que permita seleccionar la tecnología más adecuada a ser aplicada en la construcción del Sistema Inteligente de forma que responda a los requerimientos planteados.
- Seleccionar los métodos, técnicas y herramientas para evaluar un Sistema Inteligente.
- Determinar mecanismos de monitoreo y control del Sistema Inteligente una vez puesto en operación.

La consecución de los objetivos definidos en el proyecto permitirá asistir a los desarrolladores en la realización de las actividades necesarias para la implementación de Sistemas Inteligentes, basados en Machine Learning mediante un Modelo de Proceso que permita evitar comportamientos no deseados de manera que puedan crear sistemas predecibles, confiables, robustos y seguros para poder así volver a recuperar una confianza justificada de los usuarios.

Formación de Recursos Humanos

El equipo de trabajo se encuentra conformado por tres investigadores formados, tres tesistas de maestría, dos graduados de grado, un investigador de apoyo y un alumno en la carrera de Ingeniería en Sistemas de Información. Asimismo, se prevé incorporar más alumnos avanzados en la carrera de Ingeniería en Sistemas de Información con posibilidades de articular sus Trabajos Finales de Carrera de Grado, así como también tesistas de posgrado que

desarrollarán sus Trabajos Finales Integrador de Especialidad y/o Tesis de Maestría. De esta manera se espera generar un verdadero espacio integrado de investigación en el área de Sistemas de Información en el nivel de carreras de grado y posgrado.

En forma complementaria, se espera que los resultados obtenidos mejoren la comprensión por parte de los docentes integrantes del proyecto de los aspectos correspondientes a las tecnologías y métodos para la implementación de los Sistemas Inteligentes, así como, las particularidades y dificultades que tiene su aplicación en el ámbito de la educación. Esto se verá reflejado en contenidos, materiales y actividades de asignaturas de grado y posgrado.

Referencias

1. Pollo-Cattaneo, Ma. F, Pytel, P., Vegega, C., Ramón, H., Deroche, A., Straccia, L., Bernal Tomadoni, L. & Acosta, M. (2016) *Implementación de Sistemas Inteligentes para la Asistencia a Alumnos y Docentes de la Carrera de Ingeniería en Sistemas de Información*. Proceedings XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016). Workshop de Innovación en Sistemas de Software. Pág. 662-666 (Articulo 8238). ISBN 978-950-698-377-2.
2. Plawner, S., Pividori, A., Deroche, A., Vegega, C., Pytel, P., Straccia, L. & Pollo-Cattaneo, Ma. F. (2016) *Aplicación Móvil que ayude al Proceso de Enseñanza y Aprendizaje en la Asignatura Sistemas y Organizaciones de la Carrera de Ingeniería en Sistemas de Información*. Memorias IV Congreso Nacional de Ingeniería en Informática/Sistemas de Información (CoNaIISI 2016). Temática Educación en Ingeniería - Articulo ID 114. ISSN 2347-0372.
3. Pytel, P., Vegega, C., Deroche, A., Acosta, M. & Pollo-Cattaneo, MaF. (2015) *Modelo Bayesiano para el Diagnóstico del Aprendizaje en Alumnos de Inteligencia Artificial*. Workshop Tecnología Informática Aplicada en Educación (WTIAE 2015) - Articulo ID 7428. Libro de Actas XXI Congreso Argentino de Ciencias de la Computación. ISBN 978-987-3724-37-4. Premio mejor exposición del día 7 de Octubre correspondiente al XIII WTIAE.
4. Vegega, C., Deroche, A., Pytel, P., Straccia, L., Acosta, M. & Pollo-Cattaneo, Ma. F. (2016) *Diagnóstico del Proceso de Aprendizaje de Alumnos de Inteligencia Artificial mediante un Modelo Dinámico Bayesiano*. Memorias IV Congreso Nacional de Ingeniería en Informática/Sistemas de Información (CoNaIISI 2016). Temática Educación en Ingeniería - Articulo ID 19. ISSN 2347-0372.
5. Pytel, P., Vegega, C., Panizza, L., Delucchi, P., Orozco-González, M., Straccia, L. & Pollo-Cattaneo, M.F. (2018). Generación de Preguntas para Exámenes mediante la aplicación de Redes Neuronales de Aprendizaje Profundo. Memorias del 6to Congreso Nacional de Ingeniería Informática y Sistemas de Información (CONAIISI 2018). Workshop de Educación en Ingeniería. Artículo Nro 28-1274-1. ISSN 2347-0372.
6. AI Index Steering Committee (2018) *The AI Index 2018 Annual Report*. Human-Centered AI Initiative, Stanford University. https://tinyurl.com/y6whb8sl Disponible online en Febrero 2019.
7. Forbes Staff (2018). *Inversionistas están más interesados por la inteligencia artificial*. Forbes México, Portada / Tecnología https://tinyurl.com/y9mnaleu Disponible online en Febrero 2019.
8. Shah, J., Tambe, M. & Teller, A. (2016). Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel. Stanford University. https://tinyurl.com/jb3pqad Disponible online en Febrero 2019.
9. Alpaydin, E. (2016). *Machine Learning: The New AI*. MIT Press.
10. Domingos, P. (2015). *The Master Algorithm: How the Quest for the Ultimate* Learning Machine Will Remake our World. Basic Books.
11. Cohen, P. R. & Feigenbaum, E. A (2014). *The handbook of Artificial Intelligence. Vol. 3*. Butterworth-Heinemann.
12. Domingos, P. (2012). *A few useful things to know about machine learnin*g. Communications of the ACM, 55(10), 78-87. http://dl.acm.org/citation.cfm?id=2347755 Disponible online en Febrero 2019.
13. Brownlee, J. (2013). *Practical Machine Learning Problems. Start Machine Learning in Machine Learning Mastery*. https://tinyurl.com/y8nkd4l8 Disponible online en Febrero 2019.
14. Brownlee, J. (2013). A *Tour of Machine Learning Algorithms. Machine Learning Algorithms in Machine Learning Mastery*. https://tinyurl.com/y9385dsd Disponible online en Febrero 2019.
15. Walczak, S. (2001). *An empirical analysis of data requirements for financial forecasting with neural networks*. Journal of management informa□on systems, 17(4), 203-222. https://tinyurl.com/ya28wowl Disponible online en Febrero 2019.
16. Hyndman, R. J. & Kostenko, A. V. (2007). *Minimum sample size requirements for seasonal forecasting models*. Foresight, 6(Spring), 12-15. https://tinyurl.com/y8k82t3r Disponible online en
17. Collins, N. (2016, Septiembre 1). *Artificial Intelligence Will Be as Biased and Prejudiced as Its Human Creators.* Pacific Standard. https://tinyurl.com/yccjyzy2 Disponible online en Febrero 2019.
18. Crawford, K. (2016, Junio 25). *Artificial Intelligence's White Guy Problem*. The New York Times. https://tinyurl.com/y73wros8 Disponible online en Febrero 2019.
19. Büchel, B. (2018, Marzo). *Artificial intelligence could reinforce society's gender equality problems*. The Conversation. https://tinyurl.com/y74er7vt Disponible online en Febrero 2019.
20. Nuseibeh, B. & Easterbrook, S. (2000). *Requirements engineering: a roadmap*. In Proceedings of the Conference on the Future of Software Engineering (pp. 35-46). ACM.

Prototipo de Serious Game para rehabilitación de problemas neurológicos usando VRPN para comunicación entre cerebro y computadora

Javier J. Rosenstein, Rodrigo Gonzalez, Nicolás Ignacio Zárate Alvarez, Matías Campos

Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat, Belgrano 721, 5501 Godoy Cruz, Mendoza, Mendoza, Argentina.

rosensteinjavier@uch.edu.ar, gonzalezrodrigo@uch.edu.ar, nicolaszarate23@gmail.com, mgcampos9785@gmail.com

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 de ondas cerebrales. Ya sea por no disponer de un método de acceso en forma directa 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 una interfaz de realidad virtual. Este sistema debe permitir lograr el principio de neurofeedback [4], o retroalimentación hacia el paciente. De este modo se podrán mejorar sus capacidades cognitivas correspondientes. Incluso, estudios indican la posibilidad de tratar patologías psiquiátricas como la depresión [5] [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] [12].
b) Análisis de estas señales en tiempo real para poder identificar la voluntad de movimiento del individuo, así como la dirección del movimiento.
c) Transformar la voluntad de movimiento en comandos del protocolo VRPN (Virtual Reality Protocol Network) que permitan transmitir la información al componente software / hardware que la requiera.
d) Como continuación al resultado obtenido en el punto anterior, poder comandar 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] [15]. Una vez adquiridas las señales EEG/EOG, estas se analizan y codifican con las bibliotecas desarrolladas como parte de este proyecto. Esto permite su comunicación mediante VRPN con las interfaces virtuales que interpretan este protocolo.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo está compuesto por tres etapas o fases de trabajo. La primera de ellas corresponde a la implementación de VRPN entre un sistema simulado de captura de datos provenientes de un paciente, y una interfaz virtual básica. La segunda etapa del proyecto pretende avanzar sobre la captura de datos reales para que, luego de procesados, se transfieran a través de VRPN hacia la interfaz virtual de neurofeedback. Una breve descripción de cada una de las etapas se describe a continuación:

1. La primer parte consiste en la implementación de un simulador de señales 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 interfaz 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 interfaz virtual de capacitación a nivel prototipo.
2. La segunda etapa consiste en el desarrollo y mejora del proyecto mediante la adquisición real de señales EEG y EOG por electrodos ubicados superficialmente sobre la cabeza del paciente. Esta modificación al sistema requiere de un diseño e implementación electrónica así como del desarrollo del firmware que permita adquirir, analizar y procesar estas señales, que luego serán transmitidas a través del protocolo VRPN hacia una interfaz virtual. Así, el sistema diseñado se incorporará en forma transparente al

proyecto implementado en la primer etapa.

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. Se partirá 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

Los resultados esperados se pueden dividir en dos grandes grupos:

1. BCI e interfaz 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 se transmiten a través de una red Ethernet.

2. Adquisición de datos reales de EEG y EOG: análisis y preprocesamiento; se identifican 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 cuenta con la dirección del Lic. Javier J. Rosenstein (UCH) y como investigador principal el Dr. Rodrigo Gonzalez (UCH). En lo que respecta a la formación de estudiantes de la licenciatura, esta línea de investigación cuenta con dos tesis de grado en curso, pertenecientes a los estudiantes Nicolás Ignacio Zárate Alvarez y Matías Campos, cuyos planes de tesis se encuentran específicamente dentro del marco de este proyecto. Ambos cursan la licenciatura en Sistemas de Información en la Universidad Champagnat.

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. 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. Springer, 2009, pp. 137–154.
2. 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. Yathunanthan, 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, IEEE, 2008, pp. 283–288.
12. V. C. C. Roza, “Interface para tecnologia assistiva baseada em eletrooculografia,” 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. Springer, 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. IEEE, 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.

Realidad Virtual y Aumentada, Big Data y Dispositivos Móviles: Aplicaciones en Turismo

Feierherd Guillermo[1], González Federico[1], Viera Leonel[1], Romano Lucas[1], Delía Lisandro[2], Huertas Francisco[1], Depetris Beatriz[1]

[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
{gfeierherd, fgonzalez, lviera, lromano, fhuertas, bdepetris}@untdf.edu.ar

[2] III - LIDI - Facultad de Informática - Universidad Nacional de La Plata
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 30/04/2019.

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 360º.

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 el segundo año) las siguientes actividades:

1.- Identificación automática de flora mediante AR e inteligencia artificial (computer vision). La comparación realizada en el período anterior abarcaba las plataformas IBM Watson y Microsoft Azure. En esta oportunidad se agregó a las anteriores la plataforma Google Cloud, siempre buscando el reconocimiento de especies nativas de árboles del género Nothofagus.

2.- Se diseñó un algoritmo de reconocimiento que utilizaba una combinación de los resultados producidos por cada una de las plataformas en forma

individual, obteniéndose, para dos de las especies un resultado de reconocimiento del 100% (superando a todas de los que utilizaba los resultados

3. Se desarrolló una app para móviles con capacidades de realidad aumentada que invita a visitar el Patrimonio Artístico de Tierra del Fuego, con el objetivo de establecer nuevos circuitos turísticos. Durante el desarrollo se compararon diferentes técnicas para el reconocimiento de puntos de interés.

4. Se realizaron presentaciones en Congresos de la especialidad y en reuniones de investigadores.

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 que permita 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.

Ambas apps serán ofrecidas a los turistas a medida que ingresan al Parque Nacional Tierra del Fuego. Cabe señalar que a este Parque concurren el ochenta por ciento (80%) de los visitantes que arriban a la provincia por los distintos medios. La distribución está garantizada por un convenio firmado con la Administración Nacional de Parques Nacionales lo que permitirá, en caso de que resulte exitosa, extender la experiencia a otros Parques.

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), un Profesor Adjunto, un Asistente Principal JTP), un Asistente (Ayudante de 1ra.) y un Profesor Adjunto externo (UNLP), también vinculado a las TICs.

En el período de ejecución del proyecto se aprobaron una tesis de Maestría [6], una de Especialización [7] y una de grado [8]. Actualmente el Mg. González está considerando iniciar sus estudios de Doctorado en Inteligencia Artificial en la Universidad de Barcelona y el Lic. Viera la Especialización en Inteligencia de Datos orientada a Big Data en la Facultad de Informática de la Universidad Nacional de La Plata. El Lic. Romano está finalizando su Maestría en Tecnología Informática Aplicada en Educación también en la Facultad de Informática de la Universidad Nacional de La Plata.

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-aivn.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 Delía, 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-P0SJGdrM33oFEYL9Nk/edit?usp=sharing

Recomendación personalizada de documentos de Repositorio Digital Institucional aplicando minería de texto

Díaz, Fátima S., Fernández Reuter, Beatriz y Durán, Elena B.

Instituto de Investigaciones en Informática y Sistemas de Información – Facultad de Ciencias Exactas y Tecnologías – Universidad Nacional de Santiago del Estero (UNSE)

e-mail:fatimadiaz91@gmail.com; bfreuter@unse.edu.ar; eduran@unse.edu.ar

RESUMEN

Para realizar una investigación con material de calidad es fundamental contar con fuentes confiables. Es por ello que muchos recurren a repositorios digitales de universidades, dado que pueden acceder a información corroborada y utilizarla para producir nueva.

A pesar de que muchas universidades tienen sus propios repositorios, no suele sacarse el provecho adecuado de los mismos. Al existir tanta información en un solo lugar, cuesta realizar una exploración exhaustiva haciendo un uso eficiente del tiempo.

Una solución a este problema puede ser la incorporación al repositorio digital, de un sistema de recomendación personalizado, que ayude al usuario a encontrar información de manera más ágil y amigable, fomentando así el uso del repositorio.

Es por esto que se ha iniciado una línea de investigación enfocada en los sistemas de recomendación para repositorios digitales. En particular esta investigación se centra en el desarrollo de un sistema de recomendación de contenido para los usuarios del repositorio digital institucional de la Biblioteca Central de la Universidad Nacional de Santiago del Estero. Se propone la creación de un recomendador personalizado de documentos, basado en contenido, aplicando minería de texto, que contará con un perfil de usuario para que las recomendaciones puedan ser personalizadas.

Palabras clave

Sistemas de Recomendación, Repositorio Digital Institucional, Personalización, Minería de Texto

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto de investigación "Métodos y técnicas para desarrollo de aplicaciones ubicuas", código 23/C139, correspondiente a la convocatoria 2017-2020 de la Secretaría de Ciencia y Tecnología de la UNSE (SICYT – UNSE). El mencionado proyecto tiene entre sus objetivos "*Evaluar y proponer técnicas y métodos para el razonamiento y la adaptación de aplicaciones ubicuas" y "Aplicar las técnicas, métodos y estrategias propuestos en la construcción de aplicaciones ubicuas".* Con el propósito de dar cumplimiento a estos objetivos específicos, se identificaron y analizaron problemáticas locales y regionales que potencialmente pueden demandar la construcción de aplicaciones ubicuas. Entre esas problemáticas se identificó la consulta a repositorios digitales desde cualquier lugar y en cualquier momento.

1. INTRODUCCIÓN

Un *Repositorio Institucional* (RI) es un conjunto de servicios Web centralizados, creados para organizar, gestionar, preservar y ofrecer acceso libre a la producción científica, académica o de cualquier otra naturaleza cultural, en soporte digital, generada por los miembros de una institución. (De Giusti, Marisa Raquel, 2013)

En los repositorios institucionales de las universidades se suele buscar información para investigaciones científicas, puesto que la producción académica en dichas instituciones educativas avanzadas reúne resultados de investigaciones de diversas temáticas y es muy prolifera.

En esta nueva era de la digitalización las búsquedas pueden realizarse sin preocuparse por la distancia geográfica y la zona horaria. Haciendo uso de Internet podemos consultar repositorios institucionales de distintas universidades, sin tener la necesidad de trasladarnos a ellas.

Cada vez más universidades cuentan con repositorios propios, depositando en ellos diverso material de investigación, como publicaciones en congresos, revistas, capítulos de libros, tesis, tesinas, etc. A medida que el tiempo avanza, la cantidad de información va aumentando.

Realizar una búsqueda en un repositorio puede llegar a ser cansador, desmotivador y abrumador si no se tiene experiencia en búsqueda en los mismos, o incluso teniendo experiencia suele quedar la duda si la exploración realizada fue suficiente o si queda más material por explorar.

Para brindar solución a esta problemática es que surge la propuesta de prototipo de un recomendador personalizado de documentos para el repositorio digital institucional de la Universidad Nacional de Santiago del Estero.

Un *sistema recomendador* es una tecnología de filtrado de información personalizada, usada para predecir si a un usuario particular le gusta un ítem determinado (problema de predicción), o identificar un conjunto de N ítems que pueden interesarles a ciertos usuarios (problema de recomendación top-N) (Karypis, 2001).

Un sistema de recomendación *basado en contenido* se enfoca en las propiedades de los ítems. Este tipo de técnicas realizan las recomendaciones a partir de la similitud entre el contenido (información y características) de los ítems a recomendar y lo que el usuario ha señalado como de interés, lo cual permitirá personalizar la recomendación.

Existen numerosos trabajos relacionados a los sistemas de recomendación de contenido educativo y material de investigación, entre los que se pueden mencionar a Vences-Nava, Menéndez-Domínguez, & Zapata-González, (2015) quienes proponen un sistema de recomendación personalizado híbrido (combinando el filtrado colaborativo con el filtrado basado por contenido) en un repositorio de trabajos de titulación de la Universidad de Yucatán (México). Dicho sistema utiliza tanto la información contenida en los trabajos de titulación, como sus metadatos y la interacción de los usuarios, para recomendar trabajos de titulación del área de ciencias antropológicas que mejor se adapten a las necesidades de información de los usuarios.

Otro trabajo de sistema de recomendación sobre repositorios digitales, aunque en este caso sobre objetos de aprendizaje, es el realizado por Chacón-Rivas, Mario; Garita, Cesar; Solís, (2015), quienes proponen sugerir OA que podrían ser de utilidad para la lección descrita en un diseño instruccional por un profesor. Dicho sistema de recomendación emplea características propias de los sistemas

colaborativos, demográficos y de los basados en contenidos.

La propuesta de Fernández-Reuter & Durán, (2014), consiste en un buscador automático de material educativo en aulas virtuales que además recomienda, utilizando técnicas de Minería de contenido web, material adicional que tenga relación a una consulta ingresada por un estudiante. En este último trabajo se basará el sistema propuesto.

En base a los trabajos anteriormente mencionados, que forman parte de los cimientos de la propuesta, se pretende demostrar que utilizando un recomendador personalizado basado en contenido y personalizado con etiquetas de interés del usuario, se brindarían mejores sugerencias para las investigaciones realizadas, permitiendo de esta manera que el usuario encuentre los documentos que le pueden resultar de utilidad, logrando resultados más específicos en un menor tiempo.

2. DESCRIPCIÓN DE LA LÍNEA DE INVESTIGACIÓN Y DESARROLLO

En esta investigación se abordan cuestiones relacionadas a la recomendación personalizada de documentos para el Repositorio Digital Institucional de la UNSE.

Para realizar la recomendación se tendrán en cuenta los datos del título, palabras claves y resumen de las publicaciones en congreso, revista, libros, capítulos de libro, tesis, trabajos finales de grado y posgrado que se encuentren almacenados en el repositorio.

Se propone trabajar con técnicas de minería de texto para realizar la recomendación. De esta manera, se entrenará un modelo que sea capaz de reconocer las temáticas asociadas a un documento, lo que se utilizará para conformar un conjunto de documentos a recomendar al usuario. Para dicho entrenamiento, es necesario contar con la ayuda de un experto que identificará los temas con los que se relaciona cada documento del repositorio.

Se desarrollará un prototipo que permitirá determinar la validez del modelo, dado que la recomendación brindada por el sistema dependerá de la precisión del modelo en la identificación de las temáticas asociadas a la búsqueda.

A su vez, se deberá contar con un perfil de usuario con el fin de personalizar dicha recomendación, el cual contará con los temas de interés del usuario.

Como se trabajará con un repositorio ya existente, se considerarán los lineamientos, metadatos y arquitectura del RDI utilizado, de tal manera que haya una concordancia para una futura implementación.

3. OBJETIVOS Y RESULTADOS ESPERADOS

A pesar de que cada vez es más común que los sistemas acerquen al usuario aquello que puede resultar de su interés, esto no suele aplicarse en los repositorios digitales institucionales. Es por ello que, con este trabajo se espera contribuir a mejorar la difusión, búsqueda y recuperación del material almacenado en el Repositorio Digital Institucional (RDI) de la Biblioteca Central de la UNSE, recomendando automáticamente documentos personalizados según el interés de cada usuario.

Los objetivos específicos de esta investigación son:

- Desarrollar y entrenar un modelo que detecte patrones para clasificar documentos, almacenados en el RDI

de la UNSE, en base al título, resumen y palabras claves del documento.

- Diseñar y desarrollar un prototipo de recomendador personalizado, que haga uso del modelo, para brindar sugerencias personalizadas en base a las etiquetas almacenadas en el perfil del usuario, para ser implementado en el RDI de la UNSE.

Con este trabajo se espera obtener los siguientes resultados:

- Modelo de minería de texto que detecte patrones para clasificar documentos, almacenados en el repositorio digital institucional de la UNSE, en base al título, resumen y palabras claves del mismo.
- Prototipo de recomendador personalizado, que haga uso del modelo, para brindar sugerencias personalizadas en base a las etiquetas almacenadas en el perfil del usuario, para ser implementado en el RDI de la UNSE.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de esta línea de investigación se está formando una Becaria de Grado con Beca de Investigación para Estudiantes, otorgada por la Consejo de Investigación Científica y Tecnológica de la Universidad Nacional de Santiago del Estero (CICyT-UNSE), períodos 2018 y 2019, y se está desarrollando un Trabajo Final de Grado de la carrera Licenciatura en Sistemas de Información de la UNSE.

5. BIBLIOGRAFÍA

Chacón-Rivas, Mario; Garita, Cesar; Solís, J. (2015). *Agente Híbrido Recomendador de Objetos de Aprendizaje*. *Laclo*, *5*. Retrieved from https://www.researchgate.net/publication/270341748_Agente_Hibrido_Recomendador_de_Objetos_de_Aprendizaje

De Giusti, Marisa Raquel (2013). *¿Por qué conviene construir un Repositorio Institucional?*. Reunión del Consejo Interuniversitario Nacional (CIN). Retrieved from http://sedici.unlp.edu.ar/handle/10915/31682

Fernández-Reuter, B., & Durán, E. B. (2014). *Buscador automático de material educativo en aulas virtuales*. IX Congreso de Tecnología En Educación y Educación En Tecnología (TE&ET 2014), 166–173. Retrieved from http://sedici.unlp.edu.ar/handle/10915/38501

Karypis, G. (2001). *Evaluation of Item-Based Top-N Recommendation Algorithms*, 1–13. Retrieved from https://dl.acm.org/citation.cfm?id=502627

Vences-Nava, R., Menéndez-Domínguez, V., & Zapata-González, A. (2015). *Modelo híbrido de recomendación de trabajos de titulación*, (January 2016). Retrieved from https://www.researchgate.net/publication/291333144_Modelo_hibrido_de_recomendacion_de_trabajos_de_titulacion

Segmento Terreno Para Misiones Espaciales de Próxima Generación

Pablo Soligo y Jorge S. Ierache
Universidad Nacional de La Matanza, psoligo@unlam.edu.ar, jierache@unlam.edu.ar

RESUMEN

Las líneas de investigación y desarrollo presentadas tienen por objetivo demostrar la factibilidad de desarrollar sistemas de segmento terreno satelital costo-efectivos (Del inglés cost-effective) utilizando exclusivamente componentes dentro de los denominados, de estantería, en cualquiera de sus variantes OTS (Del Ingles Off-the-Shelf), COTS (Del Inglés Commercial-Off-The-Shelf) y OSS (Del Ingles Open Source Software) prescindiendo de soluciones propias o de herramientas de escasa penetración en la industria de software de propósito general.

Palabras clave: Segmento Terreno, Software, Costo-Efectivo.

CONTEXTO

La UNLaM como socio académico de la CONAE dicta una maestría en desarrollos informáticos de aplicación espacial (MDIAE). Las experiencias realizadas en la maestría, tanto de manera directa, operando unidades de software de segmento terreno de varias agencias, como mediante investigación general de las soluciones implementadas en el área, propiciaron la creación de un grupo de investigación en el marco de proyecto "Proince C-211". Este grupo está dedicado a investigar e implementar prototipos de software alternativos de bajo costo basados en las soluciones ampliamente aceptadas, de probada madurez y con penetración en la industria de software de propósito general.

INTRODUCCIÓN

Aunque no es una regla absoluta y muchas veces el alcance depende de la agencia u organización, la industria especial muestra un marcado conservadurismo en las soluciones implementadas [1], [2]. Esta realidad es definida en parte por el alto costo de las misiones y la dificultad (muchas veces imposibilidad) de realizar ajustes o reparaciones, condición que motiva la baja predisposición a tomar riesgos.

El ciclo de vida de una misión espacial está fuertemente guiado por el estricto cumplimiento de los requerimientos (requirement-driven) y esta estrategia dificulta una visión más amplia y de largo plazo [2]. No es difícil encontrar que cada agencia espacial implementa sus propias soluciones *ad-hoc*, normalmente basadas en sus protocolos y en algunos casos utilizando herramientas de escasa penetración o con años de obsolescencia.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Las líneas de investigación se proponen la exploración de propuestas alternativas basadas exclusivamente en técnicas y herramientas de uso masivo en la industria del software, en particular aplicadas a la industria espacial. El objetivo primario que se persigue es demostrar que las herramientas utilizadas en la industria del software de propósito general son aplicables a la industria espacial, puntualmente, objeto de esta investigación, el segmento terreno. Entendiendo también que la aplicación de estas herramientas colabora en el desarrollo de soluciones de bajo costo y alto nivel de mantenibilidad. Se destacan para el segmento terreno tres áreas principales de trabajo:

Lenguajes de comando y control.

Se propone el uso de lenguajes e intérpretes de propósito general para la decodificación de telemetría y el desarrollo de scripts de comandos, marginando las soluciones propias de uso común en

la industria espacial, como los lenguajes STOL , PLUTO (ESA), lenguajes de propósito especifico y normalmente incompatibles entre sí [3] La propuesta tiene como argumento, no solo la estandarización o la posibilidad de compartir procedimientos, los problemas derivados de los cambios de plataforma y los esfuerzos en las adaptaciones quedan parcialmente cubiertos por la comunidad u organización responsable del lenguaje/interprete. Esto último colabora con la reducción de costos y, en algunos casos, permite que los intérpretes puedan ejecutarse en plataformas actualizadas [4] Los recursos disponibles y el tiempo de entrenamiento en herramientas propietarias también es un costo que considerar. Los recursos disponibles y el tiempo de entrenamiento en herramientas propietarias también es un costo que considerar.

Persistencia, recuperación y explotación de datos.

Se ha optado por la utilización de un motor de base de datos relacional para el almacenamiento y recuperación tanto de los datos de la misión como de su definición. Los sistemas de bases de datos sean relacionales o no, se han transformado en una herramienta de uso extendido abarcando prácticamente todas las áreas de la industria.
Desde sencillas base de datos embebidas como Sqlite hasta complejos RDBMS como SqlServer u Oracle. En esta área, si bien existen trabajos y experiencias documentadas [5], [6] y su uso es recomendado [7], todavía existen agencias que deciden almacenar sus datos en archivos con formatos propietarios dificultando el control de integridad, el acceso, la explotación y la compatibilidad entre aplicaciones. Se exploran así las alternativas de bajo costo disponibles, los diseños y las opciones de optimización de los motores (RDBMS) que ofrezcan un rendimiento acorde a los requerimientos de un sistema informático moderno.
Castigando el rendimiento, todo el acceso a los datos se realiza mediante un ORM (Object-relational mapping), en virtud de mantener las aplicaciones tan independientes del motor de la base de datos (RDBMS) como sea posible y colaborar con ya mencionada reducción de costos.

Interfaces y distribución de procesos

En términos de interfaces tanto internas como externas en el área espacial se pueden encontrar desde soluciones propietarias hasta la implementación de tecnologías de escasa aceptación como CORBA [8]. Si bien CORBA es un estándar propuesto por varios jugadores importante de la industria del software, no ha sido adoptado masivamente y se vieron relegadas por implementaciones basadas en HTTP (Hypertext Transfer Protocol) [9].
Las interfaces del sistema propuesto se basan completamente en esta última tecnología. Toda ingestión de telemetría se realiza mediante un servicio REST que puede ser consumido tanto por módulos de terceros como por módulos internos que adapten formatos, tecnologías o características de antiguas implementaciones. De manera consistente la publicación de telemetría y datos en general se ofrece también utilizando HTTP lo que ha permitido la integración con herramientas como el NASA OPENMCT [10] (https://nasa.github.io/openmct/). (Véase Ilustración 1 - Datos de la misión argentina).

Ilustración 1 - Datos de la misión argentina SAC-D accedidos mediante el NASA OPENMCT.

Resultados obtenidos/esperados

El FS2017 [11] fue el primer segmento de vuelo utilizado como prueba experimental para el desarrollo de los sistemas planteados en las diferentes líneas de investigación. Durante esta primera integración se estableció como objetivo

desarrollar un sistema multimisión pero prescindiendo de lenguajes propietarios para la extracción y decodificación de la telemetría y utilizando motores de base de datos para la persistencia y posterior procesamiento. Utilizando un intérprete de propósito general (Python) y las capacidades de reflexión de este, se logró decodificar completamente la telemetría del modelo de ingeniería sin que esto supusiera un costo computacional relevante [12] (Ver Ilustración 2 - Tiempos de decodificación para paquetes de 5000 variables de telemetría.)

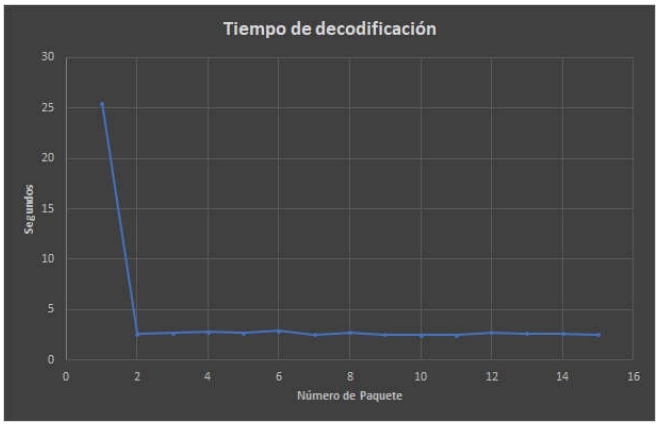


Ilustración 2 - Tiempos de decodificación para paquetes de 5000 variables de telemetría

Durante el 2018 se incorporó telemetría proveniente de la red satnogs de las misiones Lituanicasat 2 y Bugsat-1 demostrando el carácter multimisión del sistema. En el 2019 se está incorporando la telemetría histórica de la misión argentina SAC-D, siendo esta la primera misión con telemetría histórica almacenada incorporada al sistema lo que supone un desafío en términos de recuperación de datos en función del volumen. Con esta nueva incorporación se espera demostrar la capacidad de los medios de persistencia utilizados (RDBMS) para manejar de manera eficiente y segura los datos de plataforma de una misión espacial científica completa.

FORMACIÓN DE RECURSOS HUMANOS

Los trabajos desarrollados presentan una plataforma de experimentación realista con los sistemas espaciales. Permite a investigadores y alumnos probar soluciones, comprobar límites, comparar alternativas y establecer criterios de decisión. La posibilidad de trabajar con varias misiones, desde pequeñas misiones universitarias hasta datos de satélites científicos permite responder de manera concreta a las premisas de desarrollar un sistema transparente al satélite en órbita.

Junto con las líneas de investigación presentadas en este documento se desarrolla una estación terrena que permite la descarga de datos de satélites en tiempo real y la formación de recursos en los aspectos de comunicaciones y mecánica orbital. La interacción con satélites actualmente en órbita permite validar los desarrollos de dinámica orbital. Actualmente el grupo se encuentra formado por un investigador formado, tres investigadores en formación.

AGRADECIMIENTOS

Se agradece al DIIT (Departamento de Ingeniería e Investigaciones Tecnológicas) de la UNLaM por el soporte en las investigaciones. Se hace extensivo el agradecimiento al Instituto Mario Gulich y a CONAE por el acceso a la telemetría de la misión argentina SAC-D.

ACRÓNIMOS

CONAE
Comisión Nacional de Actividades Espaciales 3
CORBA
Common Object Request Broker Architecture 2
COTS
Commercial-Off-The-Shelf 1
DIIT
Departamento de Ingeniería e Investigaciones Técnologicas..... 3
MDIAE
Maestríaen Desarrollos Informáticos de Aplicación Espacial 1
ORM
Object-relational mapping 2
OSS
Open Source Software 1
OTS
Off-the-Shelf 1
RDBMS
Relational database management system 2
SAC-D
Satélite de Aplicaciones Científicas - D 3
STOL
Satellite Test and Operation Language 2
UNLaM
Universidad Nacional de La Matanza 3

BIBLIOGRAFÍA

[1] M. Petković, R. Boumghar, M. Breskvar, S. Džeroski, D. Kocev, J. Levatić, L. Lucas, A. Osojnik, B. Ženko y N. Simidjievski, «Machine learning for predicting thermal power consumption of the Mars Express Spacecraft,» *arXiv preprint arXiv:1809.00542,* 2018.

[2] J. J. Ramos Pérez, «A design for an advanced architecture of satellite ground segments,» 2014.

[3] G. Chaudhri, J. Cater y B. Kizzort, «A model for a spacecraft operations language,» de *SpaceOps 2006 Conference*, 2006.

[4] G. Garcia, «Use of Python as a Satellite Operations and Testing Automation Language,» de *GSAW2008 Conference, Redondo Beach, California*, 2008.

[5] J. Houser y M. Pecchioli, «Database Administration for Spacecraft Operations-The Integral Experience,» *ESA BULLETIN,* pp. 100-107, 2000.

[6] P. Cruce, B. Roberts, M. Bester y T. Quinn, «A database centered approach to satellite engineering data storage, access, and display,» 2007.

[7] K. Galal y R. P. Hogan, «Satellite Mission Operations Best Practices,» 2001.

[8] L. Foti, «CORBA Technology for Ground Segment System Software Development,» de *DASIA 98-Data Systems in Aerospace*, 1998.

[9] M. Henning, «The rise and fall of CORBA,» *Queue,* vol. 4, pp. 28-34, 2006.

[10] J. Trimble y G. Rinker, «Open Source Next Generation Visualization Software for Interplanetary Missions,» de *14th International Conference on Space Operations*, 2016.

[11] P. Soligo, E. González, E. Sufán, E. Arias, R. Barbieri, P. Estrada, A. Montilla, J. Robin, J. Uranga, M. C. Valenti y others, «Misión CubeSat FS2017: Desarrollo de Software para una Misión Satelital Universitaria,» *WICC 2017,* p. 843.

[12] P. Soligo y J. S. Ierache, «Software de segmento terreno de próxima generación,» *CACIC 2018.*

Sistema de gestión administrativa utilizando blockchain

Roldán, M.[1], Romagnano, M.[1,2]

[1]Instituto de Informática, FCEFN, Universidad Nacional de San Juan
[2]Departamento de Informática, FCEFN, Universidad Nacional de San Juan

maxir.unsj@gmail.com, maritaroma@iinfo.unsj.edu.ar

RESUMEN

En la actualidad, en la mayoría de las actividades diarias y rutinarias que desarrolla un ser humano interviene la tecnología. Sin embargo, aún quedan actividades donde la misma no está presente y en los tiempos que corren ésta es crucial para el seguro y correcto funcionamiento de dichas tareas. Así por ejemplo, la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan, que si bien posee un ágil y moderno sistema online para realizar las gestiones académicas, carece de uno que permita realizar trámites administrativos en los que se requiera presentar y certificar antecedentes.
En la última década se han ido popularizando los términos blockchain y bitcoin. Así por ejemplo numerosas empresas, organizaciones, instituciones gubernamentales e incluso las educativas, están haciendo uso de esta emergente tecnología para llevar a cabo sus transacciones a diario.
Las aplicaciones basadas en blockchain están aportando especificaciones, las cuales hasta hace poco tiempo eran impensables para los sistemas de gestión administrativa y académica, tales como la permanencia, transparencia e inmutabilidad. Estas características permiten que estos sistemas sean ideales para el tipo de necesidades planteadas en la facultad.
Este trabajo presenta un sistema que gestiona el procesamiento automático de antecedentes personales.

Palabras clave: administración, blockchain, educación, gestión, sistema.

CONTEXTO

Proponer un sistema que tramite y concentre información unificada y actualizada de antecedentes personales se considera un aporte importante para una institución que necesita manipular información de forma precisa y segura. En la generalidad de los casos se vuelve una tarea burocrática y compleja en cuanto a tiempo y principalmente porque puede llevar a la duplicación y/o pérdida de información, entre otras cuestiones relevantes.

1. INTRODUCCIÓN

Desde el surgimiento, en el año 2008, el bitcoin ha sido masivamente aceptado e incorporado por distintas áreas en casi todo el mundo. Sin embargo, la tecnología que lo hace posible es la clave de su gran éxito. La blockchain o cadena de bloques, es el protocolo que hace posible el intercambio de criptomonedas y que día a día se implemente en mayor número y distintas aplicaciones.
Con el paso de los años, luego de utilizar este protocolo en el ámbito de las finanzas electrónicas, la blockchain se ha implementado en otros ámbitos. De esta manera hoy en día, ésta tecnología emergente se encuentra incursionando en el área de la educación.
Sony Global Education (SGE) y Sony Corporation anunciaron el desarrollo de un sistema que aplicará específicamente la tecnología blockchain al sector educativo. Este sistema confiable centraliza la gestión de datos de múltiples instituciones educativas y permite registrar y referenciar datos educativos y transcripciones digitales. El sistema está basado en IBM Blockchain, que se entrega a

través de IBM Cloud y con la tecnología de Hyperledger Fabric 1.0, un framework blockchain y uno de los proyectos Hyperledger alojados por The Linux Foundation. En 2018, Sony implementó sus propias ofertas de servicios, comenzando con su Global Math Challenge que reúne a 150.000 participantes de todo el mundo (Sony Global Education team, 2017).

Otra empresa, Attores, ha lanzado un producto llamado Open Certificates, que permite emitir certificados educativos sobre el blockchain de Ethereum a través de contratos inteligentes. El producto está siendo probado actualmente, no es de código abierto y posee alianzas con organizaciones educativas en Singapur (Grech y Camilleri, pág. 59, 2017).

La Universidad de Nicosia de Chipre (UNIC) es una de las instituciones educativas pioneras en el área de la blockchain. Acepta Bitcoin como pago para la matrícula de cualquier programa de grado en la Universidad desde octubre de 2013. Además, ofrece un programa de grado académico acreditado (Maestría en Ciencias en Moneda Digital), que tuvo sus primeros estudiantes graduados en junio de 2016. También emite certificados académicos en la cadena de bloques de Bitcoin, utilizando su propia plataforma de software interna desde septiembre de 2014 (Grech y Camilleri, pág. 68, 2017).

Cuando se utiliza la tecnología blockchain en la emisión de certificados, se pueden verificar las credenciales sin un intermediario. El objetivo notarial de los certificados en una blockchain es, por lo tanto, convertir un certificado digital, que un estudiante normalmente recibe de forma privada, en información verificable automáticamente que puede ser consultada por terceros a través de un sistema de prueba inmutable, es decir en un blockchain público.

En la práctica actual, el acceso a una plataforma pública, casi inevitablemente, requiere que el estudiante comparta o divulgue metadatos “sensibles”, que tienden a incluir información privada. Aglietti propone utilizar una blockchain como “prueba de conocimiento”, dicha información privada no necesariamente debe divulgarse durante una consulta pública de metadatos relacionados con las certificaciones. En el corto plazo, es probable que los estudiantes puedan acercarse a institutos académicos y empleadores, mientras mantienen un discreto nivel de confidencialidad, en principio solo la información que los estudiantes marcarían como pública durante el proceso de generación de pruebas sería accesible para terceros (Aglietti, 2017).

El Instituto Tecnológico de Massachusetts (MIT Media Lab) y el Learning Machine Group (LMG), un proveedor de software empresarial, han desarrollado el estándar abierto de Blockcerts para emitir y verificar credenciales en el blockchain de Bitcoin. Blockcerts es un estándar de código abierto para emitir y verificar registros en el blockchain. El objetivo de la comunidad de Blockcerts es promover su adopción como el estándar global principal (en términos de adopción social) para emitir registros en el blockchain (MIT Media Lab, 2016).

La República de Malta ha estado considerando un programa piloto nacional en Blockchain en educación desde 2016. En enero de 2017, el Ministerio de Educación y Empleo (MEDE) firmó un Memorando de Entendimiento con Learning Machine Group (LMG). Se hizo hincapié en la intención de ambas partes de desarrollar e implementar una prueba piloto en Malta de la plataforma de tecnología de estado nación de LMG, que se basa en el estándar abierto de Blockcerts desarrollado por LMG y el MIT Media Lab. El 22 de septiembre de 2017, el Ministerio de Educación y Empleo de Malta firmó un contrato con LMG para implementar cuatro pruebas por separado en el Colegio de Malta para la Ciencia y la Tecnología de las Artes (MCAST), permitiéndoles diseñar plantillas de diploma, aprobar listas de destinatarios y emitir diplomas digitales para los graduados. Al Instituto de Estudios del Turismo (ITS), se le permitirá conferir certificados digitales de finalización a los graduados, mientras que la Comisión Nacional de Educación Superior (NCFHE) podrá crear plantillas y emitir declaraciones de equivalencia educativa para los alumnos que lo soliciten, para reemplazar

el proceso basado en PDF que se utiliza actualmente. Los beneficios de estos registros basados en blockchain son protección contra fraudes para NCFHE y verificación instantánea para cualquier entidad que desee verificar la autenticidad de la declaración. La Universidad de Malta también está en conversaciones sobre una asociación estratégica con LMG en un piloto de credencialización blockchain e investigación académica (Grech y Camilleri, pág. 74, 2017).

Desde el 2 de abril hasta el 6 de junio de 2018, aproximadamente 30 estudiantes de la Universidad de Stanford estudiaron el diseño de software de Ethereum, Quantum y NEO, así como también la arquitectura de las aplicaciones descentralizadas o DApps. Entre los proyectos desarrollados en la clase de Kokkalis se encuentra una aplicación para determinar la falsedad o veracidad de una noticia llamada Fake News Detector, DApp que permite a la comunidad votar y definir la credibilidad de artículos periodísticos a través de contratos inteligentes Solidity en la blockchain de Ethereum. Otro de los proyectos desarrollados por los estudiantes consistió en un mercado web para automatizar tareas menores y hacer más eficiente el flujo de trabajo, conocido como Web Workflow. En total se desarrollaron 16 aplicaciones basadas en Ethereum en menos de tres meses. La Universidad Stanford, ubicada en California, Estados Unidos, es una de las instituciones educativas que ha apostado por la enseñanza en temas de vanguardia como la tecnología blockchain. Stanford ofrece cursos, clases y talleres acerca de Bitcoin y desarrollos vinculados a la blockchain. De hecho, como universidad ha participado en debates sobre la escalabilidad de Bitcoin, junto al Instituto Tecnológico de Massachusetts (Rivero, 2018).

A partir del 20 de julio de 2018, la Universidad de Cagliari (UniCa) en Italia, comenzó el registro de los títulos de sus egresados en la blockchain de Ethereum, con la finalidad de garantizar su validez e integridad. María Del Zompo, rectora de la Universidad de Cagliari expresó: “Hemos decidido garantizar la autenticidad de los certificados europeos de nuestros graduados con esta tecnología, porque tenemos la intención de equiparlos con una herramienta moderna, simple e inmediata que se puede utilizar en cualquier parte del mundo. Un certificado de grado, tanto en papel como en formato digital, es fácilmente falsificable o modificable. Gracias a esta tecnología, nuestros estudiantes pueden garantizar la autenticidad e integridad de sus certificados digitales a los posibles empleadores de todo el mundo, de manera sencilla y gratuita” (Aresu, 2018).

Otro ejemplo de certificación digital en una cadena de bloques en Estados Unidos tiene lugar en el Central New Mexico Community College donde, con la ayuda de un proveedor llamado Learning Machine, desarrollaron una cartera digital que permite almacenar los certificados registrados en la blockchain. Con este sistema aspiran alcanzar la cifra de 4 mil certificados emitidos para el año 2018 (Chatlani, 2018).

En el Reino Unido también surgió una iniciativa basada en esta tecnología, llamada Woolf University, que fue promovida por un grupo de investigadores y profesores de la Universidad de Oxford. En este caso se va más allá de la certificación de los cursos, ya que el proyecto busca que el registro de toda la data administrativa de profesores y alumnos quede asentado en la cadena de bloques. Buscan crear una Universidad en línea, donde todo el trabajo administrativo utilice blockchain y así los profesores solo utilicen su tiempo para tareas educativas e investigación (Woolf University, 2018).

En India también existe un proyecto gubernamental, denominado IndiaChain, el cual, en colaboración con IIT-Bombay (Indian Institute of Technology Bombay), está haciendo un esfuerzo similar al de MIT para emitir certificados digitales a los estudiantes que se gradúan de la Universidad de Delhi a partir de 2019 (Díaz, 2018).

La Universidad de Alicante, España, anunció que realizará las primeras jornadas técnicas del Laboratorio Blockchain aplicado a las Administraciones Públicas y a las Empresas (BAES). El encuentro tiene un doble propósito. El primero es debatir sobre el alcance que tienen este tipo de redes en la

sociedad actual, el segundo es probar una blockchain de 4 nodos, de las cuatro universidades públicas valencianas, para validar la participación y la emisión de los certificados de los asistentes. Ramón Martínez, encargado de implementar el proyecto e investigador de BAES comentó: "A cada persona que fue registrándose se le asignó un token. En la entrada al evento, los tokens se devolverán a la organización vía código QR. En ese momento se genera un certificado de asistencia en formato pdf con una firma digital cualificada longeva, el cual se guardará en la red blockchain como activo y se le transferirá a cada asistente". El BAES está conformado por juristas, economistas e informáticos, con la participación de científicos del Laboratorio Europeo de Física de Partículas (CERN). Su misión es desarrollar este tipo de sistemas para el beneficio de la administración pública o privada y alcanzar un estándar europeo (Gómez, 2018).

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Una blockchain o cadena de bloques es esencialmente una base de datos distribuida de registros o un libro de contabilidad público de todas las transacciones o eventos digitales que se han ejecutado y compartido entre las partes que han participado. Cada transacción en el libro de contabilidad se verifica por consenso de la mayoría de los participantes en el sistema. Y, una vez ingresada, la información nunca puede ser borrada. La cadena de bloques contiene un registro determinado y verificable de cada transacción realizada.

Bitcoin es el ejemplo más popular que está intrínsecamente ligado a la tecnología blockchain. También es el más controvertido, ya que se ha creado un mercado global multimillonario de transacciones anónimas sin ningún control gubernamental. Por lo tanto, tiene que tratar una serie de problemas regulatorios que involucran a los gobiernos nacionales e instituciones financieras (Nachiappan, Crosby, Pattanayak, Verma, & Kalyanaraman, pág. 3, 2015).

La blockchain es almacenada por todos aquellos nodos de la red que se mantienen en sincronía con ésta. Cada bloque perteneciente a la blockchain contiene información referente a las transacciones relativas a un periodo (agrupadas en una estructura denominada Merkle Tree), la dirección criptográfica (apuntador hash) del bloque anterior y un número arbitrario único (nonce) (CriptoNoticias, 2016). Este procedimiento puede observarse en la Figura 1.

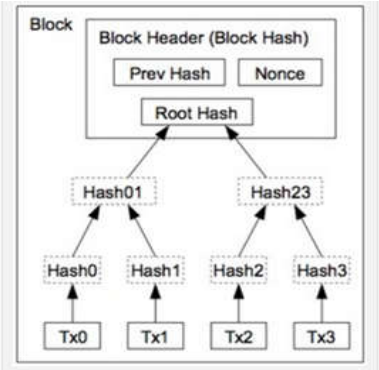


Figura 1. Estructura e información contenida en un bloque de la cadena de bloques (Fuente: Bitcoin White Paper)

El desarrollo consiste en un sistema web que permita implementar las características de la tecnología blockchain de Ethereum y contratos inteligentes para resolver distintas tareas administrativas realizadas por distintos usuarios, en la Facultad de Ciencias Exactas, Físicas y Naturales.

El proceso de investigación se llevará a cabo en cinco fases de trabajo:

Fase I: Revisión exhaustiva de antecedentes teóricos y de implementación, con el propósito de familiarizarse con las estructuras y funcionalidades de las existentes blockchains.

Fase II: Análisis de posibles problemas a resolver con la implementación de la blockchain en el ámbito educativo.

Fase III: Análisis de requisitos e implementación. Para esto se utilizarán: JavaScript, Vue.js, Node.js, Firebase Firestore, NoSQL y Ethereum.

Fase IV: Obtención de datos, pruebas y validación del sistema.

Fase V: Análisis de los resultados obtenidos, mediante una simulación, con el objeto de concluir la implementación y documentación del sistema.

3. OBJETIVOS ESPERADOS

Como objetivo general se espera desarrollar un sistema de gestión administrativa que optimice procesos de la FCEFN, utilizando la tecnología blockchain.

Además, para cumplir con dicho objetivo, se plantean los siguientes objetivos específicos:

- Identificar, analizar y evaluar los procesos que pueden mejorarse por medio del uso de tecnología blockchain.
- Especificar los procesos identificados para ser mejorados.
- Incorporar los procesos, identificados previamente, al sistema.

4. FORMACIÓN DE RECURSOS HUMANOS

Este trabajo se encuentra enmarcado dentro de la temática planteada para desarrollar la tesis de la carrera de Licenciatura en Ciencias de la Computación, de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan, San Juan.
Contribuye a su desarrollo el Instituto de Informática donde el alumno Roldán realiza sus tareas de investigación, bajo la dirección de la Mg. Romagnano.

5. BIBLIOGRAFÍA

Aglietti, A. (2017). GROWBIT @ International Open Recognition Day. Disponible en: https://tail.aqquadro.it/growbit-international-open-recognition-day-a39281072a6c.

Aresu, V. (2018). Certificati di laurea, autenticità garantita con la Blockchain. Disponible en: https://www.unica.it/unica/it/news_notizie_s1.page?contentId=NTZ121111.

Chatlani, S. (2018). Institutional Innovation: How blockchain could transform student ROI. Disponible en: https://www.educationdive.com/news/institutional-innovation-how-blockchain-could-transform-student-roi/521540/

CriptoNoticias (2016). ¿Qué es una cadena de bloques (block chain)? Disponible en: https://www.criptonoticias.com/informacion/que-es-una-cadena-de-bloques-block-chain/

Díaz G. (2018). Universidad de Cagliari certificará títulos en la blockchain de Ethereum. Disponible en: https://www.criptonoticias.com/educacion/unversidad-cagliari-certificara-titulos-blockchain-ethereum/

Gómez R. (2018). Universidades españolas certificarán jornada científica con una blockchain. Disponible en: https://www.criptonoticias.com/educacion/universidades-espanolas-certificaran-jornada-cientifica-blockchain/

Grech, A. & Camilleri, A. (2017). Blockchain in Education. Disponible en: http://publications.jrc.ec.europa.eu/repository/handle/JRC108255

MIT Media Lab (2016). What we learned from designing an academic certificates system on the Blockchain. Disponible en: https://medium.com/mit-media-lab/what-we-learned-from-designing-an-academic-certificates-system-on-the-Blockchain-34ba5874f196

Nachiappan, Crosby, M., Pattanayak, P., Verma, S., & Kalyanaraman, V. (2015). Blockchain Technology. Sutardja Center for Entrepreneurship & Technology. Disponible en: https://scet.berkeley.edu/wp-content/uploads/BlockchainPaper.pdf

Rivero J. (2018). Estudiantes de Universidad Stanford desarrollan 16 aplicaciones blockchain en menos de 3 meses. Disponible en: https://www.criptonoticias.com/educacion/estudiantes-universidad-stanford-desarrollan-16-aplicaciones-blockchain-menos-3-meses/

Sony Global Education team (2017). Sony develops system for authentication, sharing, and rights management using blockchain technology. Disponible en: https://www.sonyged.com/2017/08/10/news/press-blockchain/

Woolf University (2018). Building the first blockchain university. Disponible en: https://woolf.university/#/

Soluciones IoT con tecnología LoRaWAN

Javier Díaz, Laura Fava, Diego Vilches Antao, Agustín Candia
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, dvilches, acandia}@linti.unlp.edu.ar

RESUMEN

Para 2020, se espera que 50 mil millones de dispositivos (cosas) estén conectados a Internet para formar el mundo de Internet de las Cosas o IoT (Internet of Thing).. En el mundo de las redes de comunicación, se utilizan dos categorías principales de redes para IoT: redes de bajo consumo y corto alcance y redes bajo consumo y de largo alcance. Se considera que IPv6 sobre redes de área personal inalámbricas de baja potencia (6LoWPAN) es una red crucial en redes de corto alcance de baja potencia donde las motes de 6LoWPAN representarán la mayoría de las cosas de bajo alcance de baja potencia. Otras opciones no basadas en IP y de amplia difusión en el mercado incluye las tecnologías ZigBee y Z-wave, sobre todo en los entornos de domótica. Por otro lado, LoRaWAN y SigFox son dos de los principales entornos de redes y actores en redes de largo alcance y baja potencia o, a menudo, llamadas redes de área amplia de baja potencia (LPWAN) (Kashoash, H, Kemp, 2018). Aparecen en escena también, los protocolos asociados a la infraestructura de las redes celulares como el LTE-M y el NB-IOT.

Al mismo tiempo, el mundo se está urbanizando rápidamente y ello está provocando importantes cambios en nuestros estilos de vida, nuestro comportamiento social y nuestra salud. Un modelo de ciudad inteligente y sostenible deberá ofrecer soluciones integrando nuevos dispositivos y servicios y organizando los procesos dentro de la ciudad, promoviendo el uso eficiente de los recursos y un funcionamiento armónico del desarrollo social.

Así como la ciudad, el campo también está siendo alcanzado por nuevas tecnologías afines, que permitirán un mayor y mejor control sobre los procesos productivos.

En este contexto, se plantean experiencias de innovación tanto para las grandes ciudades como para áreas suburbanas o rurales.

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 escenarios de smart cities, áreas rurales, marítimas o alejadas de lo urbano donde la conectividad 3G o 4G resulta ineficaz/inaccesible y se presenta a LoRaWAN como red de soporte.

Palabras claves: Internet de las Cosas (IoT), sensores, protocolos para IoT, Lora, LoRaWAN.

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, estacionamiento inteligentes, horticultura, deporte, etc., desde hace varios años.

Uno de los primeros proyecto de IoT vinculado a ciudades inteligentes es *rParking: un sistema de plazas de estacionamiento*

reservadas (Boccalari, E., González, F., 2016) actualmente en funcionamiento en la playa de estacionamiento del edificio de Rectorado de la UNLP.

También se está trabajando en proyectos vinculados a IoT y deporte donde se han diseñado y construído dispositivos basados en LEDs y sensores de proximidad para mejorar la toma de decisiones, la reacción física y cognitiva, la capacidad de atención y concentración de los jugadores. En esta línea también se han diseñado y construido dispositivos basados en GPS (Global Positioning System) y acelerómetro, para medir parámetros como máximas velocidades, impactos, distancia recorridas, etc. en situación de juego (Fava, L., Vilches Antao, D., 2018).

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 el desarrollo de un kit de bastón y anteojos con sensores que detectan y avisan ante objetos en espacios bajos y altos para personas ciegas.

Las líneas de trabajo que se describirán a continuación 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.

Los proyectos antes mencionados, dan cuenta del trabajo que se viene desarrollado con IoT en diferentes escenarios, mayormente utilizando wi-fi y bluetooth, sin embargo, cuando se necesitan comunicaciones sobre distancias largas esos protocolos no son adecuados. En este sentido, han surgido nuevas tecnologías que permiten comunicación eficiente en distancias muy largas, donde no existe conectividad celular (GPRs, 3G o 4G), y también en escenarios donde las capacidades de cómputo y energía del nodo sean claramente limitadas. Dentro de estas tecnologías conocidas como LPWAN tenemos LoraWAN y Sigfox. En este artículo se presentan desarrollos de proyectos que utilizan LoraWAN para brindar soluciones a problemáticas de las ciudades y al agro.

En este artículo se describen líneas de investigación y desarrollo que forman parte de un proyecto integral destinado a proveer soluciones de IoT aplicadas a diferentes contextos locales usando LORA, como protocolo comunicación de largo alcance.

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 soluciones para ciudades inteligentes y soluciones para el agro utilizando el protocolo LoRaWAN.

Los ejes principales de I+D+i son:

- Análisis de protocolos de comunicación de largo alcance y bajo consumo. Despliegue de la red, instalación y prueba.
- Análisis de tecnologías de vanguardia para la construcción de dispositivos inteligentes para smart cities. En especial se está trabajando en la construcción de luminarias

y semáforos inteligentes para mejorar el tránsito de la ciudad.

- Diseño y construcción de dispositivos basados en sensores para el monitoreo de ganado y análisis de la salud de los animales.
- Análisis de sistemas de posicionamiento para tracking de animales: GPS y localización sin GPS usando LoRaWAN.
- Diseño e implementación de plataformas horizontales que gerencien los dispositivos mencionados, escalen en cantidad y diversidad de nodos y permitan trabajar con los datos recolectados.

3. RESULTADOS OBTENIDOS /ESPERADOS

Para la línea de IoT planteada en este artículo se ha trabajado por un lado en la red de soporte LoRaWAN y por otro en la creación de dispositivos inteligentes que serán conectados utilizando esa red y aplicaciones de gerenciamiento desarrolladas específicamente para cada tipo de dispositivo.

En cuanto a los proyectos que se están implementando podemos mencionar:

- *Iluminación inteligente*

En una primera instancia se diseñó un conjunto de dispositivos con tecnología LED, sensores de presencia, luminosidad y de calidad de aire, montados sobre una red mesh ZigBee. Equipos perimetrales que permiten dar aviso del paso de peatones y encender anticipadamente el equipo. A su vez, sensores de CO2 y material particulado, informaban en tiempo real indicadores sobre la calidad de aire. La Fig. 1 muestra el prototipo de las primeras luminarias que sirvieron de base para las nuevas investigaciones.

Fig 1.: primer prototipo de luminarias

En el contexto de un convenio de trabajo con fabricantes de luminarias led locales, se está trabajando actualmente sobre un artefacto de alumbrado público comercial dotado con sensores para gestionar su estado usando la red LoRaWAN. Se agregaron también sensores de temperatura, humedad y material particulado.

Como resultado de este trabajo, se podrá conocer el estado de funcionamiento de las luminarias en tiempo real y poder telegestionar las mismas con acciones como encendido, apagado y dimerizado. Un despliegue de este piloto ya puede observarse en el Edificio de Modernización de la Ciudad de La Plata.

- *Tracking de Ganado*

La sensorización de animales, la recolección de datos y su análisis es otra línea de trabajo. En este marco se está desarrollando un primer prototipo de collar para monitorear el ganado, detectar enfermedades y controlar posibles epidemias.

Se propone desarrollar un kit para monitorear la salud de la vaca y su ubicación en tiempo real, esto es, sensar la temperatura corporal del animal, registrar la rumia y otros parámetros que permitan determinar la salud de los hatos y la ubicación dentro del campo.

- *Semaforización Inteligente*

En el contexto de dos trabajos finales y en cooperación con un proveedor del rubro semaforización, se lleva a cabo el desarrollo de un prototipo funcional de red de semáforos inteligentes. Se incorporan funcionalidades tales como monitoreo del estado lumínico del artefacto, sincronización de onda verde, sensores de posicionamiento e inclinación, habilitación de canales para circulación ante situaciones de emergencia, entre otras. La solución pensada de extremo a extremo, incluye el diseño de un nodo controlador, facilidades de comunicación vía LoRaWAN, monitoreo en tiempo real y herramientas de configuración y analitica. La Fig. 2 muestra el prototipo base de semáforo que se está usando en esta primera etapa.

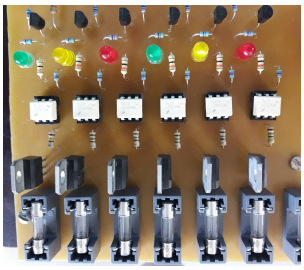


Fig 2.: Prototipo de semáforo base

- *Sensores de calidad de aire*

Se implementó un nodo de monitoreo de calidad de aire, basado en el uso de sensores portátiles. Se analizó la performance de 3 modelos diferentes de sensores para la medición de material particulado, uno de los contaminantes primarios determinado por la Organización Mundial de la Salud (OMS, 2006).

Como parte del análisis exploratorio, se realizaron campañas de monitoreo en conjunto con un equipo de referencia certificado, a fines de analizar desvíos en las mediciones y correlación entre los diferentes sensores. La tarea se desarrolló en conjunto con profesionales del ámbito de la tecnología ambiental pertenecientes a la UNLP (Candia, A. , 2018).

- *Balizamiento en Boyas*

Como parte de un proyecto que permita dar soporte a un sistema de balizamiento en Puertos, se diseñó un nodo a batería para ser instalado en el km 10.4 sobre el Río de La Plata, partido de Ensenada. El objetivo fue medir el desempeño de las comunicaciones LoRaWAN sobre espejos de agua. El resultado obtenido fue más que satisfactorio, logrando captar paquetes de la mayoría de los equipos de comunicación instalados en la ciudad de La Plata.

Finalmente, cabe destacar que para dar soporte a todos estos proyectos se desplegó una red LoRaWAN usando la infraestructura de torres de la UNLP y el backbone de fibra óptica que las conecta. La arquitectura se implementa en una topología de estrellas en la que dispositivos repetidores (gateways) retransmiten mensajes entre los dispositivos finales (sensores) y un servidor de red central

Los gateways están conectados al servidor de red a través de conexiones IP estándar y actúan simplemente convirtiendo paquetes de RF en paquetes IP y viceversa. El despliegue de gateways en las antenas seleccionadas, cubrió de forma satisfactoria el área de interés, incluso utilizando una baja densidad de equipos. La Fig. 3 ilustra los equipos desplegados y el área cubierta en las pruebas de alcance realizadas.

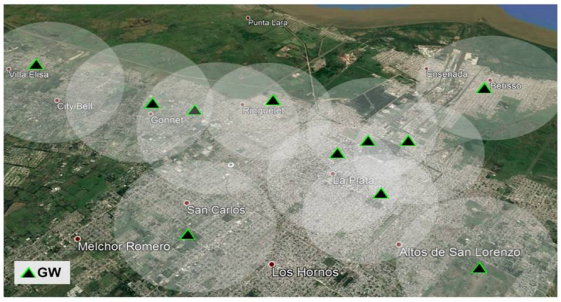


Fig 3.: Mapa de cobertura LoRaWAN. La Plata, Berisso y Ensenada

4. Formación de Recursos Humanos

El equipo de trabajo de la línea de I+D+i presentada en esta 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 la Facultad de Informática y a la Facultad de Ingeniería.

En relación a las tesinas de grado vinculadas con esta línea de investigación, se está dirigiendo una tesina de grado y una PPS vinculadas a semaforización inteligente. Se está comenzando con una tesina de dos alumnos de las licenciaturas para trabajar con IoT y Agro, y una tesina con IoT y seguridad. También se están ejecutando cuatro proyectos acreditados de la Secretaría de Políticas Universitarias del Ministerio de Educación y Deportes de la Nación vinculados a medición de calidad de aire, agro, deporte, accesibilidad y luminarias donde participan docentes del LINTI y alumnos de las distintas carreras de las facultad mencionadas.

5. Referencias

Boccalari, E., González, F. *rParking: Sistema de plazas de estacionamiento reservadas.* Tesina de Grado accesible en http://sedici.unlp.edu.ar/handle/10915/59485. Octubre, 2016.

Fava, L., Vilches Antao, D., Díaz, J., Pagano, M., Romero Dapozo, R., *Tecnología aplicada al deporte de alto rendimiento,* XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), ISBN: 978-987-3619-27-4. Abril, 2018.

Kashoash, H, Kemp, A. *Comparison of 6LoWPAN and LPWAN for the Internet of Things.* Australian Journal of Electrical and Electronics Engineering. Diciembre, 2017.

Candia, A., Luengo, M. Represa, S., Porta, A. Marrone, L. *Solutions for SmartCities: proposal of a monitoring* system of air quality based on a LoRaWAN network with low-cost sensors. ISBN: 978-1-5386-5447-7, 2018.

OMS, «Guia de calidad del aire,» 2006.

Recursos/Desarrollos Existentes

Ciudades Sustentables, *https://www.argentina.gob.ar/ambiente/ciudadessustentables*

LORA Alliance, *https://www.lora-alliance.org*

Design SupportTechnical resources and documents, *https://www.semtech.com/design-support*

Una propuesta de decisión multicriterio en IdeTR

Sandra Rodriguez[1] , Emilio Rearte[2], Marta Cabrera Villafañe[3], Elena Pauli[4], Mariela Fernández[5], Víctor Hugo Loyola[6], Marcela Analía Melidoro[7],Sandra Oviedo[8], Daniel Díaz Araya[9]

[1,2,3,4,5,6,7] Centro de Investigación y Desarrollo Informático (CIDI) / Sede Universitaria Chamical / Universidad Nacional de la Rioja.
[8,9]Laboratorio de Informática Aplicada a la Innovación, Instituto de Informática/ Facultad de Ciencias Exactas, Físicas y Naturales /
Universidad Nacional de San Juan

sandraorona27@yahoo.com.ar, emiliorearte@yahoo.com.ar, marta.cabreravilla@gmail.com, elenapauli@hotmail.com, mariel.lucia@gmail.com, loyolavh@gmail.com, lic.marcelamelidoro@gmail.com, soviedo@iinfo.unsj.edu.ar, ddiaz@iinfo.unsj.edu.ar

Resumen

La evaluación y selección de ideas de un portfolio de ideas es una tarea que requiere de un método objetivo que permita disminuir las subjetividades de quienes tienen esta responsabilidad. Diferentes intereses entran en juego, muchas veces contrapuestos, lo que puede llevar a cometer graves errores cuando se trata de elegir una idea con el interés de desarrollar una innovación. La subjetividad se puede controlar mediante el uso de Metodos de Desición Multicriterios. La plataforma Ideas en Tiempo Real (IdeTR) es un desarrollo del Laboratorio de Informática Aplicada a la Innovacion de la UNSJ, que pretende asistir al proceso innovador, ya cuenta con el asistente para el proceso generador de ideas, se espera avanzar con el prototipado de la herramienta que asista a la evaluación y selección de ideas.

Palabras clave:

Métodos de decisión multicriterio. Toma de decisiones. Innovación. Creatividad. Evaluación de ideas.

Introducción

Los Métodos de Decisión Multicriterio (MDM), surgen en el ámbito empresarial con el propósito de facilitar la toma de decisiones en situaciones complejas en las que existen alternativamente, más de un curso de resolución posible. Con el tiempo se convirtieron en una herramienta muy potente para solucionar problemas de selección en general, pero de aplicación especialmente útil en contextos de decisión definidos por varios elementos y variables, así como de posiciones o intereses contrapuestos.

Ideas en Tiempo Real (IdeTR) es una Plataforma Tecnológica para la Generación de Ideas de Productos Tecnológicos, 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. IdeTR tiene un componente que asiste durante el proceso de creatividad de manera online, como resultado se obtiene un conjunto o portfolio de ideas [1]. Cuando se desea prototipar o avanzar con la ejecución de una idea, en una etapa siguiente del proceso innovador y hay que elegir una idea, diferentes intereses, en muchos casos contrapuestos surgen, dificultando la tarea.

Esta investigación propone desarrollar una alternativa metodológica para la toma de decisiones en los procesos de selección de ideas aplicando principios de las teorías de decisión multicriterio. Esta búsqueda permitirá reducir riesgos en la toma de decisiones, hasta hoy, basada solamente en la intuición y la experiencia de quien evalúa las ideas. Se repasarán los principales métodos de decisión multicriterio para la toma de decisiones, se evaluarán los criterios de selección y se definirán métricas de ponderación en pos a disminuir las consideraciones humanas (subjetivas) y construir a una selección más objetiva. Se espera tener como resultado el prototipo un componente para sumar completitud a IdeTR.

Contexto

La línea de investigación se desarrolla en el contexto del Centro de Investigación y Desarrollo Informático (CIDI) de la Sede Universitaria Chamical de la Universidad Nacional de la Rioja, bajo el proyecto de investigación denominado "Una propuesta de decisión multicriterio para la evaluación y selección de ideas innovadoras en IdeTR". Este proyecto se encuentra aprobado por el programa de incentivos en colaboración con 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.

Líneas de Investigación

1. Toma de Decisiones

Se puede definir a la toma de decisiones como "Un proceso intencional que combina el análisis de la información, la confrontación de alternativas, la valoración de las opciones y, finalmente, la toma de la decisión". Es condición para un problema de decisión multicriterio la presencia de más de un criterio y que estos se encuentren en conflicto. Es decir, un problema puede considerarse como problema multicriterio si y solo si existen al menos dos criterios en conflicto y al menos dos alternativas de solución, para definir los atributos que se utilizan para describir alternativas que puedan ser cuantitativas (objetivas) o cualitativas (subjetivas) y una meta u objetivo (González & Garza, 2003). Esto lleva a la necesidad de

establecer el uso de métodos sistemáticos, es decir, estructurados y coherentes en el campo disciplinar, técnico o académico de la decisión; elegir entre las posibilidades, previamente identificadas, la mejor forma de proceder valorando las implicaciones y consecuencias de tal forma que se revele el nivel de compromiso con ella. [2].

2. Un enfoque holístico para la evaluación de ideas innovadoras

Innovación es una palabra muy usada actualmente por las empresas que desean mejorar su competitividad a través de la creatividad y así fomentar la capacidad de innovación. La mayoría de las empresas utilizan sus propios recursos para estas actividades, lo que conlleva a una demanda de calidad en procesos cortos de tiempos pero que a la vez sean efectivos para generar, evaluar y explotar ideas. [3]. Una idea innovadora se refiere específicamente a un producto de tecnología potencialmente convertible en un negocio sustentable; esto es, económicamente rentable, social y ambientalmente beneficioso, así como sostenible en el tiempo [4]. Pues bien, ante esta necesidad de evaluar y seleccionar una idea innovadora de un conjunto de ideas resultantes de una sesión de creatividad es una tarea compleja. Se destaca el valor de las ideas, donde una mala evaluación puede llevar a descartar ideas con mucho potencial, provocando grandes pérdidas. Por esta razón, precisamente, uno de los requerimientos indispensables del entorno de trabajo en que se desarrolla esta actividad consiste en la posibilidad de brindar una participación activa y colaborativa de todos los participantes con el objetivo de aceptar y sumar diferentes perspectivas de la realidad social y del mercado, considerar todas las visiones acerca del rol que debe ocupar la tecnología en este contexto, valorar simultáneamente múltiples dimensiones del producto y, finalmente, disminuir la subjetividad de los aportes.

Centrados en las empresas, la principal preocupación no es solo la selección de ideas, sino también si las ideas seleccionadas podrían contribuir a la competitividad sostenida y crear un nivel de beneficio suficiente [3].

Debe tenerse en cuenta que el desarrollo de ideas innovadoras incluye el establecimiento de objetivos, la identificación de esos objetivos y la evaluación de capacidades que la empresa tiene a su disposición, como puede verse en la figura 1.

3. Proceso para el desarrollo de ideas innovadoras

Para comprender la complejidad de todo el proceso, desde la identificación de las ideas hasta su implementación, se debe descomponer en fases consecutivas. Las etapas son: 1. Identificación y organización de ideas 2. Hacer una lista de los métodos y

técnicas disponibles 3. Crear un conjunto de criterios para seleccionar un método o herramienta en particular 4. Seleccionar la herramienta o el método 5. Implementar el método 6. Seleccionando la idea 7. Implementación de la idea, como se detalla en la figura 2 [3].

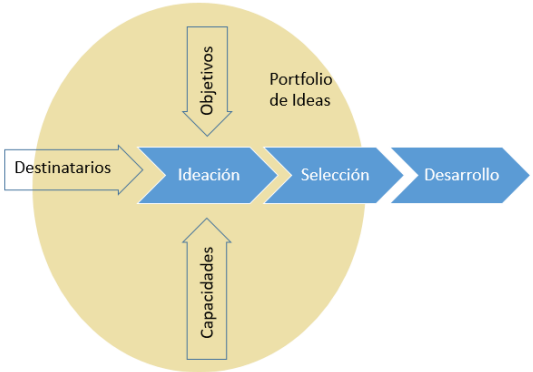


Figura 1 Estructura del portfolio de ideas, basado en [3].

Hay que notar que este proceso se puede aplicar una vez que se han concebido las ideas, es decir que se cuenta con un portfolio de ideas resultantes, por ejemplo de una sesión de creatividad.

Una vez identificados los criterios generales de evaluación y de selección de ideas a adoptar y desarrollar, se elegirá el método y modelo a adoptar en cada caso. Los métodos de evaluación valoran una serie de factores que tratarán de garantizar el éxito de la idea. Estos factores pueden ser diferentes criterios de evaluación, por ejemplo, tecnología empleada, posible rentabilidad, grado de novedad de la idea, consumo de energía, impacto ambiental, características de usabilidad, entre otros, muchas veces estos criterios de evaluación resultan contrapuestos. Para cada idea que haya superado en primera instancia los criterios generales, se le aplicará la evaluación numérica de los factores identificados. Estos factores habrán sido ponderados previamente. Finalmente se aplicará el algoritmo correspondiente al modelo de selección mas adecuado [5].

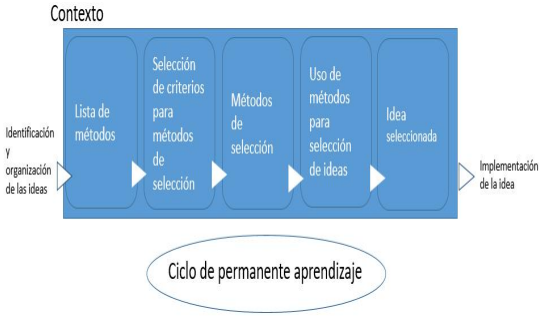


Figura 2 Proceso desde la identificación de ideas hasta su implementación.

4. El proceso de toma de decisiones multicriterio

En respuesta a los requerimientos apuntados para el tratamiento de problemas de selección se han desarrollado procedimientos de análisis denominados como de Métodos de Decisión Multicriterio (Multiple Criteria Decision Making o MCDM), también reconocidos como Métodos de Decisión Multiobjetivos. La metodología de los Métodos Multicriterio consiste en determinar criterios de evaluación asignándoles los pesos respectivos, los cuales se analizan a través de comparaciones pareadas para medir la relación entre ellos y su importancia.

Posteriormente, expertos evalúan las alternativas por cada criterio, asignan una calificación y obtienen así la alternativa óptima para el problema en cuestión. [6].

Formación de Recursos Humanos

El equipo de trabajo que lleva adelante este proyecto se compone de 8 docentes investigadores de diferentes sedes universitarias, 2 tesistas de grado en período iniciación, 1 tesista de posgrado (maestría) iniciando su trabajo.

Conclusiones

Los MDM son herramientas para manejar la subjetividad humana, la toma de decisiones resulta un proceso complejo y con un cierto grado de incertidumbre que se desarrolla conforme a determinados criterios, valores y preferencias. El desarrollo actual de la metodología multicriterio es una de las ramas o vertientes considerada como columna básica de la investigación operativa, [6].

Cada método tiene sus propias características y son convenientes para su aplicación en diferentes tipos de situaciones de decisión.

El proyecto se encuentra en desarrollo y actualmente se está trabajando en el estudio de las principales áreas temáticas, variables, alternativas y criterios que serán empleados y que permitirán analizar, describir aquellos métodos que más se adapten a la solución propuesta.

En este trabajo se desea establecer pautas y criterios de compatibilidad conceptual y procedimental que permitan acoplar la herramienta a la plataforma IdeTR del Laboratorio de Informática aplicada a la innovación de la Universidad Nacional de San Juan.

Referencias

[1] Diaz,, Oviedo, & Otazú, Ideas en tiempo real, un asistente para la generación de ideas en tiempo real, JATIC 2017. UCAECE, Mar del Plata.

[2]Artieta Pinedo & González Labra, La toma de decisiones 1998, Madrid. Ed. Trotta, pp. 367-407.

[3] Rebernik & Bradač -Idea evaluation methods and techniques. Creative Treiner. Union Europea.

[4] Prim, 3 Pasos para Encontrar Ideas Innovadoras para crear una Empresa; ;https://innokabi.com/como-generar-ideas-innovadoras-tormenta-de-ideas/

[5].http://www.eoi.es/wiki/index.php/An%C3%A1lisis_y_selecci%C3%B3n_de_ideas_en_Competitividad

[6] Chica-Salgado, C. A. (2013). Modelo matemático multicriterio para coadyuvar a la toma de decisiones en la selección de alternativas en Pymes. Estrategias, 11(21), 49-63

Procesamiento de Señales y Sistemas de Tiempo Real

Detección inteligente de cambios y generación de alarmas tempranas en sistemas de monitoreo

Ricardo Coppo[1], Patricia Baldini[1], Héctor Bambill[1], Marina Jakomin[2], Guillermo Calandrini[3], Javier Balmaceda[1], Andrés Fernandez[1], Santiago Tarnoski[1], María Tourret[1], Camila Silva Bustos y Julián Abdala[1].

[1]Grupo de I+D en Sistemas y Tecnologías de la Información y las Comunicaciones (Grupo SiTIC).
Dpto. de Ingeniería Electrónica, Facultad Regional Bahía Blanca UTN, 11 de Abril 461, Bahía Blanca

[2]Dirección de cuencas, Dirección Nacional de Política Hídrica y Coordinación Federal, Secretaría de Infraestructura y Política Hídrica. Ministerio del Interior. Esmeralda 255, piso 11, CABA.

[3]Dpto. de Ingeniería Eléctrica y de Computadoras, Universidad Nacional del Sur, San Andrés 800, Bahía Blanca.

{rcoppo, pnbaldi, hbambill}@frbb.utn.edu.ar; mjakobin@yahoo.com; calandri@criba.edu.ar

RESUMEN

Un cambio o anomalía en un sistema de monitoreo se refiere al instante en que se produce una variación abrupta significativa de alguno de sus parámetros con respecto al patrón habitual de comportamiento. La detección de estos eventos son utilizadas en sistemas de alerta temprana (SAT) meteorológicas, ambientales, en mantenimiento de maquinarias, y el monitoreo de servidores informáticos, entre otros. Para la detección de cambios de estado se analizan los errores que resultan de la comparación entre los valores obtenidos mediante un modelo predictivo con los medidos por la red de monitoreo. Variaciones significativas de este error pueden ser atribuidas a cambios que disparen una alerta. El modelo puede ser estadístico, basado en el procesamiento adaptativo de señales y, en forma mas reciente, en la inteligencia computacional.
En este proyecto se propone en una primera fase, el estudio de diferentes enfoques en el planteo de modelos predictivos y detectores de cambio orientados en particular al monitoreo del medio ambiente y de entornos industriales. En una segunda fase se plantea la evaluación e implementación de algoritmos de detección de cambio aplicados tanto a señales de origen simulado como provenientes de prototipos experimentales de laboratorio.

Palabras Clave: detección de eventos; sistemas de alerta temprana; procesamiento adaptativo; inteligencia computacional.

CONTEXTO

La línea de investigación presentada está en ejecución desde enero de 2019, en el marco del proyecto acreditado y financiado por la Universidad Tecnológica Nacional titulado: "Detección inteligente de cambios y generación de alarmas tempranas en sistemas de monitoreo".
Su ámbito de realización es el Grupo SiTIC (Grupo de I+D en Sistemas y Tecnologías de la Información y las Comunicaciones) Facultad Regional Bahía Blanca de la UTN. Se cuenta, además con la participación de un investigador del Departamento de Ingeniería Eléctrica y de Computadoras de la Universidad Nacional del Sur y con la asistencia técnica de una profesional de la Secretaría de infraestructura y política hídrica del Ministerio del Interior.

1. INTRODUCCION

Un cambio, evento, o anomalía en un sistema de monitoreo se refiere al instante en que el mismo sufre una variación abrupta de alguno de sus parámetros de funcionamiento. Estos cambios no son necesariamente de magnitud elevada, sino que representan alteraciones significativas del estado del mismo. Estos eventos se evidencian en series de tiempo, en secuencias de imágenes, y en datos heterogéneos no estructurados y son utilizados en SAT para señales sísmicas, meteorológicas y del medio ambiente [24], el monitoreo de calidad del agua y del aire [2][7][8][25][30], el reconocimiento de voz, aplicaciones médicas [5][30][34], en sistemas financieros, en mantenimiento predictivo de maquinarias [3][22][26][33], entre otras [35]. Todos estos sistemas, aunque diferentes en la forma de captura y presentación de la información,

poseen características funcionales similares: una etapa de adquisición, una de fusión y análisis de la información adquirida, una de selección y definición de un modelo predictivo adecuado, estocástico o determinístico, con el cálculo del error de predicción (residuales) y una de generación de una señal de alarma/anomalía al determinar la presencia de un cambio de estado significativo, [10][12][14][17][21][23][28].

Los enfoques tradicionales analizan una única señal y su función auto-correlación para determinar umbrales sobre los valores muestreados o alguno de sus parámetros (amplitud, frecuencia, potencia, media, varianza). Superado este valor límite el sistema emite la señal de alarma. En entornos reales estos límites pueden ser superados ocasionalmente por la presencia de valores atípicos producidos por la existencia de ruido en los sensores, en la transmisión de datos o en el sistema de control y supervisión SCADA, disparando una falsa alarma. Si estas se producen en forma frecuente el sistema de monitoreo pierde confiabilidad. Una simple modificación de los umbrales no siempre es una solución válida del problema, ya que si se los amplia en exceso el sistema pierde sensibilidad y aumenta el error de la "no detección" de eventos reales. Por otra parte, los diversos modelos predictivos permiten la detección en base a la comparación de cada conjunto de medidas con las predichas por el modelo adoptado generando un error residual o de predicción para la observación. Si este error no presenta una distribución probabilística similar a la del ruido blanco se presupone la existencia de un cambio o anomalía. Este enfoque ha permitido una extensa investigación en los campo de los filtros digitales y teoría de control adaptativo [6][12][13][15][32].

En forma más reciente, la disponibilidad de gran cantidad de sensores distribuidos de bajo costo, configurados en redes que monitorean el entorno a frecuencias de muestreo muy altas, permiten el tratamiento del problema de detección de anomalías mediante algoritmos de *BIG DATA* o de *datamining* temporal. Con este nuevo enfoque se puede adicionar, a las series de tiempo tradicionales, información heterogénea (bases de datos con campos no numéricos, documentos de texto, imágenes, páginas web) para su procesamiento[19][20].

Por otra parte, los avances en la inteligencia computacional, reconocimiento de patrones y de la minería de datos permiten actualmente plantear modelos predictivos no lineales basados en redes neuronales, máquinas de soporte vectorial, aprendizaje profundo (*deep learning*), y modelos de grafos probabilísticos, por mencionar algunos. Estos predictores se apoyan en el reconocimiento del comportamiento del sistema observado de antemano (etapa de aprendizaje, aprendizaje supervisado) o sobre la inferencia de categorías de comportamiento similar (clusters, aprendizaje no supervisado). Los modelos obtenidos reducen el error de predicción permitiendo un análisis de inferencia de las características predominantes del sistema [3][9][16][18][19][26][27]. Los modelos de aprendizaje supervisados procesan vectores multi-paramétricos de atributos clasificados por un experto en forma previa. A partir de estos datos se construye un modelo, en la forma de una red neuronal, árbol de decisión, red Bayesiana, máquina de soporte vectorial, subespacio de estados, etc, que permite realizar la predicción para nuevas observaciones que se obtienen de la red de sensores. Un desvío significativo entre la predicción y la observación más reciente producirá la detección del evento o anomalía buscada [1].

En el aprendizaje no supervisado es innecesaria la participación de un experto previa al entrenamiento del sistema. Los algoritmos inteligentes reconocen clases o "estados" de equivalencia correspondientes a diferentes modos de operación. Al recibir una nueva observación si la clase asignada no es igual a la anterior puede interpretarse que hubo un cambio de comportamiento del sistema. Este enfoque, si bien se adapta a situaciones en las que no se conocen los posibles estados de operación del sistema, es de alta complejidad computacional y difícil de implementar en la práctica ya que requiere mantener toda la base de datos de entrenamiento a disposición del algoritmo para mejorar la robustez [1][11].

Los avances recientes en inteligencia computacional permiten formular el problema de detección de cambio en base a algoritmos de optimización con la selección de métricas adecuadas. Los conceptos de velocidad de detección del cambio, el tiempo medio entre falsas alarmas, la exactitud de la determinación del tiempo de cambio, y la sensibilidad general del algoritmo constituyen parámetros deseables de cuantificar y comparar.

En este proyecto se propone el planteo experimental de diferentes modelos predictivos y de algoritmos de detección de puntos de cambios. Se trabajará sobre series de tiempo simuladas y obtenidas de diversos sensores reales instalados sobre una plataforma de ensayo a fin de evaluar las características,

bondades y limitaciones de los mismos. Para finalizar, se debe señalar que pasar de la problemática de la detección de una anomalía a la generación de una alerta temprana y la decisión de poner en marcha un mecanismo de respuesta, posee consecuencias que superan el ámbito del proyecto y requiere de la formación de equipos multidisciplinarios.

2. LINEAS DE INVESTIGACION Y DESARROLLO

El problema de detección de puntos de cambio o anomalías en sistemas de monitoreo para alertas tempranas requiere del aporte de varias disciplinas en la formulación del riesgo, el sistema electrónico de monitoreo, el procesamiento computacional y las eventuales acciones de mitigación del problema.

En este proyecto, de carácter fundamentalmente experimental, se busca formar recursos humanos e informáticos para abordar los aspectos computacionales en las etapas de monitoreo electrónico y de selección, implementación y evaluación de algoritmos en la detección de las fallas o anomalías.

Debido a la amplitud y a las múltiples aplicaciones de la temática, se plantea su estudio inicial en el marco de dos ejes de interés para la región de influencia de la Facultad Regional Bahía Blanca. El primero asociado al medio ambiente y el otro referido al monitoreo de condición orientado al mantenimiento predictivo de maquinarias industriales. Se prestará especial atención a series provenientes de observaciones meteorológicas e hidrológicas en lo que hace a cantidad y calidad del agua para diversos usos.

Por otra parte, la reciente instalación de un parque eólico cercano a la ciudad de Bahía Blanca ofrece la posibilidad de estudiar la conformación de SAT para mejorar los mecanismos de predicción de fallas posibles en su funcionamiento. En ambos casos se requiere de sistemas de monitoreo multisensor con capacidad para detección temprana de cambios de estado significativos.

Para cada eje se desarrollarán las siguientes actividades:

-Relevar y caracterizar los SAT, sistemas de monitoreo y el uso de algoritmos de detección de puntos de cambio. A su vez se procederá a la selección de casos factibles de simular y de probar en un banco de laboratorio adecuadamente diseñado a tal fin.

-Generación de secuencias de datos simuladas. Se usarán como series testigo de prueba y verificación para analizar los algoritmos, provenientes tanto de bases de datos de acceso público como generadas por simulación mediante programas informáticos de desarrollo propio.

-Construcción de un banco de prueba con la red de sensores y el equipamiento necesarios para estudiar problemas planteados en los ejes mencionados. Esta es la etapa constructiva de hardware más significativa del proyecto. Se registrarán mediciones en condiciones de operación "normal" y con la imposición de cambios controlados. A partir de las series observadas se analizará el efecto del ruido de los sensores, la influencia de la disposición geométrica de los mismos y las dificultades electrónicas de sincronizar el muestreo de los diferentes sensores de la red.

-Ensayo de algoritmos y comparación de bondades y limitaciones de diferentes metodologías en base a criterios adecuados. Se espera que los mayores resultados cualitativos y cuantitativos de la investigación se produzcan en esta fase.

-Generación de un banco de software modular para facilitar la construcción de diferentes modelos. Se aplicará una metodología de desarrollo de software ágil basada en el paradigma de la programación extrema para enfatizar la adaptabilidad a los cambios de las especificaciones. El desarrollo de software se realizará en la modalidad *open-source* para facilitar su divulgación. Es de interés implementar un sistema modular que permita integrar diferentes series, probar e intercambiar algoritmos en un marco unificado.

-Transferencia de conocimientos y resultados al medio socio-productivo local y regional. Además de publicar los resultados obtenidos en eventos científicos, se contempla el dictado de cursos de capacitación y la posible asistencia para entidades interesadas en instalar o mejorar sus SAT.

3. RESULTADOS ESPERADOS

Se espera poder realizar contribuciones al avance científico y/o tecnológico en varios planos debido a la gran cantidad de aplicaciones que hacen uso de los sistemas de monitoreo y alerta temprana. La posible interacción con el grupo GEIA (Grupo de Estudio en Ingeniería Ambiental) de esta Facultad Regional permitirá encontrar aplicaciones en la prevención y mitigación de eventos que podrían afectar el medio ambiente, la salud poblacional o la calidad del aire o agua de un sistema. Por otro lado, la aplicación al monitoreo de condición en entornos

industriales, en conjunto con el Departamento de Ingeniería Mecánica podría traer como resultado la implementación de algoritmos que mejoren el esquema de mantenimiento preventivo de sus maquinarias.

En una etapa futura se espera poder relacionarse con organismos no universitarios interesados en la evaluación del cambio climático, alerta frente a catástrofes naturales de diversa índoles, y mantenimiento de plantas industriales entre otros.

El proyecto tiene como objetivo principal la generación de una biblioteca de software a publicar bajo la modalidad de *open-source* para fomentar el intercambio y transferencia entre investigadores a través de la Internet. El diseño de dicho software debe contemplar una estructura modular conformada por diversos algoritmos de detección y que permita el uso de diferentes bases de datos, tanto en un modo de análisis como el de una eventual implementación en un sistema productivo. Se espera también el desarrollo de diversas actividades de extensión y divulgación en la forma de cursos y seminarios sobre los aspectos teóricos del problema y análisis descriptivo de los algoritmos del software desarrollado. Así mismo, será de interés la participación de otros organismos, públicos y privados, con necesidad de SAT o detección temprana de fallas en la forma de proyectos de colaboración interinstitucional para el estudio de casos de aplicación real.

4. FORMACIÓN DE RECURSOS HUMANOS

La intención fundamental es crear un grupo de investigación que logre especializarse en los aspectos algorítmicos del tema de detección de anomalías (modelado, procesamiento, inteligencia computacional). El equipo buscará integrarse a grupos multidisciplinarios con los especialistas correspondientes a cada eje seleccionado. Al momento de iniciar el presente proyecto se cuenta con la participación en de cinco alumnos de grado y un graduado que se inician en actividades de I+D y transferencia al medio. Como objetivo fundamental se busca que se compenetren en la metodología de trabajo asociada a las tareas de investigación científica y adquieran destreza en el uso de herramientas computacionales, desarrollo de prototipos y sensado.

Se entiende que los nuevos conocimientos sumados al enfoque interdisciplinario del proyecto y la orientación a problemas del medio socio-productivo local representan un valor agregado a su formación tradicional. La experiencia adquirida podrá definir la temática de sus respectivos proyectos finales de carrera. Además se buscará motivarlos para emprender estudios de posgrado en temáticas afines.

Al momento actual, están en proceso solicitudes de becas otorgadas por la Secretería de Asuntos Estudiantiles (SAE) y por la Secretaría de Ciencia y Tecnología de la Facultad, a fin de formalizar la participación de alumnos y el graduado. Finalmente, se espera transferir algunos resultados al medio socio-productivo, mediante cursos y seminarios de capacitación y/o divulgación.

5. BIBLIOGRAFIA

[1] S. Aminikhanghahi and D. Cook. A Survey of Methods for Time Series Change Point Detection. Knowledge and Information Systems, 51(2):339-367, May 2017.

[2] V. Z. Antonopoulos, D. Papamichail, and K. Mitsiou. Statistical and Trend Analysis of Water Quality and Quantity Data for the Strymon River in Greece. Hydrology and Earth System Sci., 5(4):679-692, 2001.

[3] A. Arcos Jimenez, M. Muñoz Gomez, and F. García Marquez. Machine Learning for Wind Turbine Blades Maintenance Management. Energies,11(1):13,Dec.2017.

[4] P. Baldini, G. Calandrini, P. Doñate and H. Bambill. PSO Algorithm-based Robust Design of PID Controller for Variable Time-delay Systems: AQM application. Journal of Computer Science and Technology, 15(2), Nov 2015.

[5] I. Barnett and J.-P. Onnela. Change Point Detection in Correlation Networks. Scientific Reports, 6(18893), Jan 2016.

[6] S. M. Bozic. Digital and Kalman Filtering. Edward Arnold Publishers Ltd. London, England, 2nd ed., 1979.

[7] R. Coppo, L. M. Jakomin y C. Delrieux. Predicción de parámetros de calidad de agua mediante redes neuronales en la cuenca del río Pilcomayo - Argentina. IV Congreso Bianual PRODECA-PROIMCA, Bahía Blanca, Sep. 2017.

[8] S. Dejus, A. Nescerecka, and T. Juhna. On-line Drinking Water Contamination Event Detection Methods. 11th Enviornment, Technology, Resources Scientific and Practical Conf., v.1, 77-81, Latvia, 2017.

[9] F. Desobry, M. Davy and C. Doncarli. AnOnline Kernel Change-Detection Algorithm. IEEE Transactions in Signal Processing, 53(8):2961-2974, Aug. 2005.

[10] R.Domingues , M.Filippone , P Michiardi.

and J. Zouaoui. A comparative evaluation of outlier detection algorithms: Experiments and analyses. Pattern Recognition,74:406–421, 2018.
[11] R. Duda, P. Hart, and D. Stork. Pattern Clasication. John Wiley and Sons, New York, 2nd ed , 2001.
[12] F. Gustafsson. Adaptive Filtering and Change Detection. John Wiley & Sons Ltd., West Sussex, England, 2000.
[13] L. B. Jackson. Digital Filters and Signal Processing. Kluwer Academic Publishers, Hingham, USA, 1986.
[14] N. A. James and D. S. Matesson. ecp: An R Package for Nonparametric Multiple Change Point Analysis of Multivariate Data. Journal of Statistical Software, Articles, 62(7):1-25, 2015.
[15] Y. Kawara, Y. Yairi, and K. Machida. Change-point Detection in Tme point Test for High-Dimensional Data. Proceedings of the National Academy of Sciences, 114(15):3873-3878, 2017.
[16] N. Khan, S. McClean, S. Zhang, and C. Nugent. Optimal Parameter Exploration for Online Change-point Detection in Activity Monitoring using Genetic Algorithms. Sensors, 16(11):17-84, 2016.
[17] D. Koller and N. Friedman. Probabilistic Graphical Models - Principles and Techniques. MIT Press, Cambridge, Massachusetts, USA, 2009.
[18] K. K. Korkas and P. Fryzlewicz. Multiple Change-point Detection for Nonstationary Time Series using Wild Binary Segmentation. Statistica Sinica, 27(1):287-311, 2017.
[19] L. I. Kuncheva and W. J. Faithful. PCA Feature Extraction for Change Detection in Multidimensional Unlabeled Data. IEEE Transactions on Neural Networks Learning Systems, 24(5):69-80, 2014.
[20] Y. Li and R. Lund. Multiple Changepoint Detection using Metadata. Journal of Climate, 28(10):4199-4216, 2015.
[21] R. Lund, X. L. Wang, Q. Lu, J. Reeves, C. Gallagher and Y. Feng. Changepoint Detection in Periodic and Autocorrelated Time Series. Journal of Climate, 20(20):5178-5190, 2007.
[22] M. Morshedizadeh. Condition Monitoring of Wind Turbines using Intelligent Machine Learning Techniques. PhD thesis, University of Windsor, 2017.
[23] H. V. Poor and O. Hadjiliadis. Quickest Detection. Cambridge University Press, Cambridge, UK, 2nd ed., 2008.
[24] J. Reeves, J. Chen, X L. Wang, R. Lund, and Q. Lu. A Review and Comparison of Changepoint Detection Techniques for Climate Data. Jour. Applied Meteorology and Climatology, 46(6):900-915, 2007.
[25] N. Sankary and A. Ostfeld. Inline Mobile Sensors for Contaminant Early Warning Enhancement in Water Distribution Systems. Jour. Water Resources Planning and Management,143(2):5401-6073, 2017.
[26] M. Schlechtingen and I. Santos. Comparative Analysis of Neural Network and Regression Based Condition Monitoring Approaches for Wind Turbine Fault Detection. Mechanical Systems and Signal Processing, 25(5):1849-1875, 2011.
[27] X. Shi, Y. Wu and C. R. Rao. Consistent and Powerful Graph-based Change. point Test for High-Dimensional Data. Proceedings of the National Academy of Sciences, 114(15):3873-3878, 2017.
[28] G. Shmueli. Practical Time Series Forecasting - A Hands-on Guide. Axelrod Schnall Publishers, USA, 3th ed., 2018.
[29] A. Tartakovsky, I. Nikiforov and M. Basseville. Sequential Analysis - Hypothesis Testing and Changepoint Detection. CRC Press - Taylor & Francis Group, Boca Raton, FL, USA, 2015.
[30] S. Tinelli and I. Juran. Numerical Modeling of Early Bio-contamination in a Water Distribution System and Comparison with Laboratory Experiments. International Conference on Sustainable Infrastructure 2017, American Society of Civil Engineers (ASCE), pages 258-269, New York, USA, Oct. 2017.
[31] C. Truong, L. Oudre and N. Vayatis. A Review of Change Point Detection Methods. The Computing Research Repository(CoRR), abs/1801.00718, 2018.
[32] R. D. Turner. Gaussian Processes for State Space Models and Change Point Detection. PhD thesis, University of Cambridge, Cambridge, UK, July 2011.
[33] A. Ukil and R. Zivanovic. The Detection of Abrupt Changes using Recursive Identication for Power System Fault Analysis. Electric Power Systems Research, 77(3):259-265, 2007.
[34] P. Yang, G. Dumount and J. Ansermino. Adaptive Change Detection in Heart Rate Trend Monitoring in Anesthetized Children. IEEE Trans. Biomedical Eng., 53(11):2211-2219, Nov. 2006.
[35] X. Xu, H. Liu and M. Yao. Recent Progress of Anomaly Detection. Complexity. 2019. https://doi.org/10.1155/2019/2686378.

Robótica, Simulación y Redes de Sensores en Sistemas de Tiempo Real

Fernando Romero[1], Diego Encinas[1], Armando De Giusti[1,2], Santiago Medina[1], Martín Pi Puig[1], Horacio Villagarcía[1,3], Juan Manuel Paniego[1], Fernando G. Tinetti[1,3]

[1]Instituto de Investigación en Informática LIDI (III-LIDI)[3]
Facultad de Informática – Universidad Nacional de La Plata – Centro Asociado CIC

[2] CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3] Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

{fromero, dencinas, degiusti, smedina, mpipuig, hvw, jmpaniego, fernando}@lidi.info.unlp.edu.ar

Resumen

El objetivo de esta línea de investigación es el estudio, desarrollo y aplicación de Sistemas de Tiempo Real (STR), en particular los que incumben a robots y redes de sensores, tanto inalámbricas como cableadas. Se llevan a cabo experimentos con diferentes tipos de robots y con plataformas basadas en microcontroladores, diseñados y armados en el laboratorio de Tiempo Real. Una sublínea importante de los trabajos la componen las comparaciones en escenarios reales y simulados.

Contexto

Esta línea de Investigación forma parte del proyecto 11/F024 – Computación de Alto Desempeño: Arquitecturas, Algoritmos, Métricas de rendimiento y Aplicaciones en HPC, Big Data, Robótica, Señales y Tiempo Real SubProyecto CAD-3. Procesamiento para problemas de Tiempo Real / Robótica del Instituto de Investigación en Informática LIDI acreditado por la UNLP.

Palabras Claves: Tiempo Real, Simulación, Sistemas Embebidos, Comunicaciones, Redes de Sensores, Robots, Microcontroladores, Cloud Computing.

1. Introducción

Los Sistemas de Tiempo Real (STR) se distinguen porque elaboran una respuesta ante una entrada determinada, pero es condición determinante que la misma debe ocurrir dentro de un plazo [5] [6] [11] [12] [16] [17]. Muchas veces son sistemas embebidos dentro de otro sistema al cual controlan, por lo que deben respetar los tiempos del sistema controlado. Esto incluye un entorno físico en donde se toman datos a través de sensores, a partir de los cuales se elaboran respuestas, que en algunos casos vuelven al entorno físico a través de actuadores. Dentro del campo de la robótica se experimenta con robot móviles, en lo que atañe a la odometría, la ubicación por GPS en exteriores y balizamiento en interiores combinados con diferentes tipos de sensores de tecnología ultrasonido e infrarrojo.

Se realizan experimentos sobre robots armados por alumnos, robots Khepera

[18], drones Parrot, placas de desarrollo basadas en microcontroladores (como, Arduino, NodeMCU, CIAA [19] [27]) y Computadoras de Placa Simple (como, Raspberry Pi y Banana Pi), empleándose diferentes SOTR (Linux RT-Preempt, FreeRTOS, MQX, OSEK-OS, etc.) [8].

Se analiza el funcionamiento de diferentes tipos de redes de sensores basadas en los protocolos cableados RS485, SPI, I2C, CANBUS [15] [20] y MODBUS [1]. También se realizan pruebas de alcance, integridad y funcionalidad de redes de sensores inalámbricas [28] [29] [30] principalmente utilizando modulos WiFi y LoRa [26].

Se trabaja en el campo del modelado y simulación con el fin de obtener simuladores que permitan predecir la eficiencia de distintos sistemas ante diferentes escenarios. En particular, los sistemas analizados son los siguientes: redes de sensores [31], arquitecturas de cloud computing y evacuación de edificios. Las herramientas utilizadas para los desarrollos son: Proteus [25], CloudSim [2], ABMS [14] y SystemC [24].

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 [21]. 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, eventual-mente, 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].
- Construcción de redes inalámbricas específicamente orientadas al control de variables físicas con 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. [26] [9].
- 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 con robots móviles.
- Medición de consumo energético de diferentes microcontroladores en distintas condiciones de uso y corriendo diferentes SOTR.
- Estudio de plataformas de hardware: Arduino, Intel Galileo, CIAA, Freescale Kinetis, Raspberry Pi, NodeMCU.

- Construcción y estudio de redes de sensores cableadas, empleando CANBUS, MODBUS y RS485.
- Construcción y estudio de redes de sensores inalámbricas basadas en WiFi y LoRa.
- Desarrollo de interfaces para el control de drones.
- Modelado y simulación de arquitecturas de cloud computing para comparar con arquitecturas de HPC.

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 y Analista en TICs. En cuanto a postgrado, investigadores del grupo están desarrollando un trabajo final de especialización, tres tesis de Maestría y una tesis de Doctorado.

5. Referencias

[1] Jordi Bartolomé "El protocolo MODBUS", 2011. En http://www.tolaemon.com/docs/modbus.htm

[2] 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.

[3] Azizi, F., N. Houshangi. "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] 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.
[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 PractitionerFourth 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.
[31] M. Pi Puig; S. Medina; A. Batista; D. Encinas; F. Romero; F. Tinetti; A. De Giusti. Design of a CAN Simulation Device for Communications in Sensor Networks. Computer Science & Technology Series - XXII Argentine Congress of Computer Science. Selected Papers. 2017.

Software libre. Interface gráfica para SWT

Héctor R. Lépez / Departamento de Informática / FCEFN / UNSJ / lepezhr@yahoo.com.ar

María Luisa Gordillo / Departamento de Informática / FCEFN / UNSJ / mgordillo13@gmail.com

RESUMEN

Scilab es un software gratuito y de código abierto para computación numérica, apto para ejecutarse bajo diversos sistemas operativos: Windows, Linux y Mac OS X.

Este software contiene un toolbox (caja de herramientas) "atom" (en lenguaje de Scilab), el SWT (Scilab Wavelet Toolbox) para el procesamiento de señales e imágenes a través de wavelets, el cual emula el toolbox de Matlab (MATrix LABoratory) Wavelet Toolbox™ (WT™).

Una de las diferencias que hace de WT™ una caja de herramientas más amigable con el usuario es su interface gráfica: el **wavemenu**, este es una interface gráfica para acceder a las diversas herramientas proporcionadas en WT™ de forma más rápida.

A diferencia de esto SWT, siendo tan robusto como WT™, no posee una interface gráfica.

Este trabajo es un inicio de lo que una vez finalizado queremos que sea una interface gráfica amigable de SWT, a la que hemos dado el nombre de **swavemenu.**

Palabras claves: Scilab, Wavelet, toolbox, interface gráfica.

CONTEXTO

La idea surge en el contexto del proyecto de investigación "**Análisis Matemático aplicado a Procesamiento de Señales con Wavelets, utilizando Scilab",** sin estar entre los objetivos planteados para él. Este proyecto es acreditado y financiado por CICITCA-UNSJ (2018-2019).

1. INTRODUCCIÓN

Las wavelets proveen un conjunto de herramientas maleables para abordar problemas prácticos en ciencias e ingeniería. En la actualidad, son el instrumento más poderoso para el procesamiento de señales e imágenes. La transformada wavelet resulta esencialmente eficaz para extraer información de señales no periódicas. Otra de las grandes mejorías frente a otras técnicas, es que se dispone de una amplia familia de wavelets, lo que permite su gran versatilidad para tratar señales e imágenes de diversa índole. Con ellas se puede realizar análisis (separar aproximación de detalles, identificar tendencias, puntos de corte o de cambio abrupto) y síntesis (reconstruir la señal sintetizada sin pérdida de información relevante); se puede reducir el ruido (denoising), y también comprimir la señal o imagen con escasa pérdida de información (de vital importancia tanto en la transmisión de grandes cantidades de datos, como en su almacenamiento). Debido a esto, son útiles en un amplio número de aplicaciones de muy diferentes campos, donde muchas veces se obtienen mejores resultados que con otras técnicas. Particularmente se ha comprobado, en los últimos treinta años, que las wavelets son una herramienta muy potente y con gran potencial para el trabajo con señales, en lo que

a compresión y reducción de ruido se refiere. La moderna teoría de wavelets, cuyos comienzos datan de cuatro décadas atrás, ha experimentado un gran progreso impulsado tanto desde la matemática pura y aplicada, como del dominio informático necesario para su aplicación y sustentación. Esto se debe principalmente a que el procesamiento mediante wavelets se presenta muy eficiente donde otras técnicas aplicadas antiguamente, como el análisis de Fourier, no resultan satisfactorias.

No está entre los alcances de este trabajo enseñar el sustrato matemático de la extraordinaria teoría de wavelets, ni el mostrar cómo se realiza y que resultados arroja el procesamiento de señales o imágenes hecho con SWT o WT™.

OBJETIVO/S:

El objetivo general de este trabajo es desarrollar una herramienta para SWT, a la que llamaremos **swavemenu**, que resulte una interface gráfica amigable y fácilmente comprensible por parte del usuario para el tratamiento de señales unidimensionales (que dependen de una variable independiente); una propuesta sencilla para el científico no matemático ni informático que sea tan clara y atractiva como el **wavemenu** del toolbox de WT™, pero que pertenezca a un sofware libre, en este caso Scilab. Para ello estamos trabajando con una licencia de MatLab, facilitada por el IEE (Instituto de Energía Eléctrica de la UNSJ), y con la última versión disponible de Scilab, la versión 6.0.2.

Con el propósito de alcanzar nuestro objetivo observamos cómo trabaja el wavemenu de Matlab. Éste se presenta con un esquema de dos columnas con distinto número de filas las que se agrupan bajo un mismo nombre según el tipo de tareas que ejecutan. Cada entrada o "botón" en estas columnas tiene un nombre diferente, de acuerdo a la función específica que realiza.

De todos los botones o entradas en el wavemenu, comenzamos imitando los botones:

1- "Wavelet 1-D" de los agrupados bajo el nombre One-Dimensional 1-D (wavelets discretas que proporcionan procesamiento de señales discretas uno dimensionales).

2- "SWT Denoising 1-D" (Stationary Wavelet Transform Denoising 1-D) para eliminación del ruido.

3- "Wavelet Coefficients Selection 1-D", el cual proporciona los coeficientes de la señal original y los de las señales de aproximación y de detalles obtenidos por el análisis con una determinada wavelet y un determinado nivel de descomposición.

4-"Signal Extension" que extiende las entradas de una señal a la próxima potencia de dos, para la efectiva aplicación del análisis.

Los últimos tres botones están agrupados en wavemenu bajo el nombre Specialized Tools 1-D.

El porqué de la imitación de estos y no de otros botones se debe a que recientemente hemos comenzado a incursionar en el tema de facilitar una interface gráfica al atom SWT de Scilab, y a que dentro del proyecto de investigación del cual se ha desprendido la idea de "construir" esta herramienta, declarado en el ítem CONTEXTO, sólo nos dedicamos al procesado señales discretas unidimensionales. No hemos incursionado en procesamiento de imágenes porque no ha sido objetivo propuesto en este proyecto. Por otra parte se trabaja con wavelets discretas, porque las señales en su

mayoría son series (sucesiones), es decir información digital, por lo cual el enorme esfuerzo de comprensión matemática que significaría utilizar análisis wavelet continuo (el que exige un vasto conocimiento de Análisis Matemático Superior), no reporta un beneficio substancial en comparación a los resultados obtenidos con el uso de wavelets discretas en las series (información digital) que nos han brindado para su estudio distintos investigadores.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Desarrollo de sofware libre para Procesamiento de Señales digitales utilizando Silaba.

3. RESULTADOS OBTENIDOS/ESPERADOS

En principio, se formuló la herramienta swavemenu a través de programación en Scilab, y el uso de funciones del atom SWT del mencionado software libre. Luego para estudiar el comportamiento de esta nueva herramienta y compararla con wavemenu de MatLab, respecto a cada una de sus funciones, a saber: análisis y síntesis de una señal filtrada con distintas wavelets, tendencia (aproximación), detalles, coeficientes de las subseñales, reconstrucción de la señal filtrada y las gráficas de cada una de las citadas, utilizamos dos series de distintos orígenes. La primera formulada analíticamente por $f(t) = 10t^2(2-t)^6 sen(5\pi t)$, muestreada a 1024 puntos en el intervalo [0; 2], lo que la convierte de analógica en digital, y la segunda, llamada DEJUA, obtenida de las mediciones anuales de los derrames del Río San Juan en el período desde 1909 hasta 2017. Esta última se extiende a través de la imitación de Signal Extension en swavemenu, debido a que el número de sus entradas (108) no es una potencia de dos. Es decir trabajamos con una señal "ficticia" y otra real. Las wavelets utilizadas fueron:

- (Harr, Daubechies (db N), con N desde 1 hasta 10.
- Symlet (Sym N), con N de 2 a 8.
- Coiflets (coif N) con N de 1 a 5.

Luego de probado el swavemenu en estas dos series y comparando los resultados que sobre estas ofrece wavemenu concluimos que los objetivos planteados se han cumplido considerablemente. Nos proponemos completar la emulación de estos botones con la parte estadística correspondiente y los complementos de umbralización y cuantización de aquellos que incluyen estas herramientas.

Es de esperar hacia el futuro, la incorporación de nuevos botones, que contiene wavemenu, a nuestro swavemenu, en lo que respecta al procesamiento de imágenes y a la utilización de Análisis Wavelet continuo. Esto dentro del marco de un futuro proyecto de investigación.

4. FORMACIÓN DE RECURSOS HUMANOS

A este respecto cabe destacar que el proyecto de investigación **Análisis Matemático aplicado a Procesamiento de Señales con Wavelets, utilizando Scilab,** prevé la incorporación de becarios, como así también de tesistas de grado, lo que aún no se ha producido. Esperamos pronto contar con estos recursos y que la participación en este WICC favorezca la difusión de esta temática entusiasmando a posibles becarios y/o tesistas.

5. BIBLIOGRAFÍA

1. S.Mallat, **"A Wavelet tour of Signal Processing"** Academic Press, 2nd Edition 1999. http://www.cs.nyu.edu/cs/faculty/mallat/Wavetour_fig/index.html
2. **Wavelet Toolbox™ User's Guide.**
3. **Help of Wavelet Toolbox - Scilab.**
4. **"Tratamiento de Señales Digitales mediante wavelets y su uso con Matlab"** E. Gómez-Luna- D. Silva- G. Aponte. I.S.B.N: 84-8454-387-0.
5. **"Uso de la Transformada de Ondeletas (Wavelet Transform) en la Reducción de Ruidos en las Señales de Campo Eléctrico producidas por rayos"-** Información Tecnológica Vol. 23(1), 65-78 (2012) - Francisco Santamaría, Camilo A. Cortés y Francisco J. Román. http://dx.doi.org/10.4067/S0718-07642012000100008
6. **La Transformada Wavelet Aplicada a los Sistemas Eléctricos de Potencia: Estado del Arte** (Spanish Edition) (Spanish) Paperback – August 12, 2011- Jenifer Gallego Tejada, Ferley Castro Aranda.

Procesamiento

Distribuido

y Paralelo

Aplicaciones de cómputo intensivo con impacto social: Salud y Consumo de energía en HPC

Javier Balladini[1], Marina Morán[1], Claudia Rozas[1], Candelaria Álvarez[1], Rafael Zurita[1], Rodrigo Cañibano[1], Pablo Bruno[1], Belén Casanova[1], Cristina Orlandi[2], Armando De Giusti[3], Remo Suppi[4], Dolores Rexachs[4], Emilio Luque[4]

[1] Facultad de Informática, Universidad Nacional del Comahue
{javier.balladini, marina, claudia.rozas, candelaria.alvarez, rafa, rodrigo.canibano}@fi.uncoma.edu.ar;
{psbruno90, mb.casanova.retamal}@gmail.com

[2] Hospital Francisco Lopez Lima
orlandi.mariacristina@gmail.com

[3] Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata
degiusti@lidi.info.unlp.edu.ar

[4] 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. Actualmente, las dos líneas principales de investigación que llevamos adelante son: aplicaciones para la salud con problemas de Big Data, y consumo energético de los sistemas de computación de altas prestaciones. Estas líneas de investigación se desarrollan en colaboración con una universidad nacional y otra del extranjero, y un hospital público.

Palabras claves: computación de altas prestaciones, eficiencia energética, aplicaciones para la salud.

1. Contexto

Las líneas de investigación aquí presentadas están enmarcadas 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.

Uno de los dos 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, pertenecientes al "Hospital Francisco López Lima" de la ciudad de General Roca, provincia de Río Negro. 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 HPC, ha ganado una enorme difusión en el público general debido al impedimento para seguir aumentando la frecuencia de reloj de los procesadores; el problema se encuentra fundamentalmente 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 de la arquitectura de la mayoría de los sistemas

de cómputo, incluyendo computadoras personales de escritorio, notebooks, celulares y tablets.

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 las 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, como el sector salud. 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 y comunicación entre tareas, y complejidad de la arquitectura de las plataformas hardware. A su vez, una de las mayores problemáticas que afectan a estos sistemas de cómputo de grandes dimensiones es el elevado consumo energético y potencia eléctrica demandada.

Nuestro interés está centrado en dos líneas: aplicaciones de cómputo intensivo para la salud, y gestión del 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, las Unidades de Cuidados Intensivos (UCI) proveen cuidados continuos y rigurosos para personas adultas que están críticamente enfermas y que pueden beneficiarse de tratamiento, y dan un buen morir a los pacientes irrecuperables. En las Unidades de Vigilancia Intermedia (UVI), se atienden pacientes que no requieren de tratamientos de soporte vital pero que, por su condición, son necesarios mayores cuidados y controles de los que se reciben en una unidad de internación convencional. En ambas unidades (UCI y UVI), los datos de los pacientes involucran datos clínicos de baja frecuencia, y flujos de datos fisiológicos de alta frecuencia generado por el equipamiento médico (como monitores médicos de signos vitales, respiradores artificiales, etc.). En una UCI/UVI típica, los enfermeros completan manualmente datos en formularios, registrando datos clínicos y fisiológicos. Los datos fisiológicos se obtienen por observación de las pantallas del equipamiento médico (ubicados a los lados de las camas) a intervalos de tiempo (horas o fracciones) que definen los médicos para cada paciente. Cabe mencionar que el equipamiento médico emite alertas, siempre limitadas a lecturas de sus propios parámetros (sin posibilidad de interconexión con otros equipos o de suministrarle información adicional para emitir alertas más inteligentes). Al final del día, los médicos analizan exhaustivamente los datos de los formularios y dan a los enfermeros las indicaciones de tratamientos a realizar a los pacientes. Todo este trabajo es llevado a cabo por enfermeros y médicos de manera manual, sin la asistencia de ningún sistema informático.

Los principales problemas que ocurren en las UCI y UVI típicas son:

1. La metodología actual de gestión de la información es muy propensa a errores humanos.
   Causas:
   - Podrían cometerse errores en el registro manual, que realizan los enfermeros, de datos observados en los monitores de signos vitales u otro equipamiento médico.
   - El análisis realizado por los médicos, que interrelaciona una gran cantidad de variables, provenientes de enormes volúmenes de información de diferente tipo, origen y soporte (datos fisiológicos, radiografías, análisis de laboratorio, datos de observación del paciente, etc.) es una tarea sumamente compleja. Ella, sin asistencia informática alguna, es difícil de concretar sin cometer errores.
   Efectos: Posibles diagnósticos erróneos o información inconsistente que implica un esfuerzo adicional del personal para detectar su origen.

2. Detección tardía del deterioro de la salud de los pacientes.
   Causas: Desde que los datos se toman por los enfermeros hasta que son analizados por los médicos, puede transcurrir un tiempo prolongado.
   Efectos: Incidencia negativa en la salud de los pacientes, que puede impactar en:
   - Aumento del riesgo de muerte.
   - Aumento de los costos económicos de los tratamientos.
   - Estadías más largas, y por ende la atención de menos pacientes.

3. Pérdida de información de pacientes.
   Causas: Los datos observados en el equipamiento médico es registrado en formularios a intervalos de horas (o varios minutos), perdiendo todas las muestras intermedias (por ejemplo, el monitor de signos vitales actualiza la frecuencia cardíaca una vez por segundo).
   Efectos:

- Disminución de la precisión de los diagnósticos.
- Imposibilidad de registrar el conjunto completo de todos los datos medidos por el equipamiento médico para uso futuro. Disponer de un gran conjunto de datos históricos permite la extracción de nuevo conocimiento, útil para futuras investigaciones médicas y el descubrimiento de patrones para predicción del patologías. La predicción de patologías, como sepsis (una de las principales causas de muerte en UCIs), permitiría comenzar los tratamientos más temprano, reduciendo el riesgo de muerte, el costo económico de los tratamientos, y las estadías de los pacientes.

Un gran avance sería disponer de un sistema de cómputo que detecte el deterioro en la salud de los pacientes en tiempo real, basándose en múltiples parámetros de diferentes orígenes. La detección temprana del deterioro permitirá aumentar la efectividad de los tratamientos, y por consiguiente la mejora de la salud de los pacientes, la reducción del costo económico de los tratamientos, y el aumento del número de pacientes atendidos. El problema principal que debe enfrentarse para la construcción de este sistema, y que creemos viable con la aplicación de técnicas de computación paralela y distribuida, 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 [11, 3, 1, 2].

Consumo energético de sistemas de HPC

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 [12, 10].

El consumo energético es hoy en día un gran problema. Para dar una idea de la magnitud del mismo, la máquina de mayor prestaciones de la actualidad, la máquina de EEUU Summit, demanda casi 10 MW de potencia, lo mismo que se requiere para abastecer a los hogares de una ciudad con alrededor de 135.000 habitantes (calculo 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 pocos cores rápidos), uso adecuado de la jerarquía de memoria, hibernación de recursos, escalado dinámico de frecuencia y tensión, rediseño de algoritmos, planificación de tareas, y 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, y el consumo energético de los sistemas de HPC.

Aplicaciones para la salud

Esta línea está enfocada en el diseño y desarrollo de un sistema para detección automática y temprana del deterioro de la salud de pacientes en Unidades de Terapia Intensiva (UTI) y 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.

El sistema está basado en reglas clínicas. Cada regla define las condiciones, relacionando parámetros y valores, que deben cumplirse para ge-

nerar una alerta indicando posible riesgo (actual o futuro) en la salud de los pacientes. Cada paciente puede asociarse a un conjunto de reglas particulares, creadas desde el conocimiento de expertos o guías clínicas. El sistema adquiere datos desde sistemas de historia clínica electrónica y del equipamiento médico como los monitores de signos vitales. Las reglas clínicas asociadas a cada paciente son continuamente evaluadas (en tiempo real) y, si una de ellas concuerda, se emite una alerta.

La construcción del sistema involucra diferentes desafíos tecnológicos relacionados fundamentalmente a:

- Extracción de datos del equipamiento médico.

- Procesamiento en tiempo real de grandes volúmenes de datos (Big Data).

Consumo energético de sistemas de HPC

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 erroneos.

- Diseñar la infraestructura de un sistema de Big Data de tiempo real para el procesamiento de reglas médicas.

- Diseñar y desarrollar una aplicación para la interacción del sistema con médicos y enfermeros.

Los avances/resultados actualmente comprenden:

- Un análisis del estado general de las UTI del hospital Francisco Lopez Lima, y la propuesta del diseño de alto nivel del sistema [8].

- El diseño de la infraestructura del sistema de procesamiento de reglas y un prototipo, y el desarrollo de un dispositivo embebido adquisidor de datos (actualmente obtiene información de la señal analógica del electrocardiograma desde un monitor médico) que envía datos a la plataforma de procesamiento por medio de una red de WiFi [4].

- Estudios para la realización de un dispositivo para obtener la saturación de oxígeno en sangre de los pacientes, y su integración con el dispositivo adquisidor de datos.

- El desarrollo parcial de una aplicación que detecta complejos QRS de manera eficiente en señales de electrocardiogramas.

Consumo energético de sistemas de HPC

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) [7], 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, y en [6] expusimos un análisis de los factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. Actualmente, hemos agregado la predicción para la operación de restart (adicionalmente a la de checkpoint), y se contemplan distintas alternativas de configuraciones del sistema, relacionadas a: almacenamiento en NFS, configuración de la aplicación de checkpoint/restart (compresión/sin compresión), 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 local a la Universidad Nacional del Comahue, cuenta con un doctor en la temática (recibido en 2008). En Marzo de 2019 se ha presentado la tesis de grado "Una Infraestructura de Big Data de Tiempo Real para un Sistema de Detección del Deterioro de Pacientes en Terapia Intensiva", en la Universidad Nacional del Comahue. En el transcurso del 2019 se espera la finalización de dos tesis de grado en temas de aplicaciones para la salud, y en 2020 una tesis doctoral en el tema de consumo energético de sistemas de HPC.

Referencias

[1] Amara health analytics, http://www.amarahealthanalytics.com (accedido en marzo de 2019).

[2] ehcos smarticu, http://www.ehcos.com/productos/ehcos-smarticu/ (accedido en marzo de 2019).

[3] Excel medical, http://excel-medical.com/ (accedido en marzo de 2019).

[4] Javier Balladini, Pablo Bruno, Rafael Zurita, and Cristina Orlandi. An automatic and early detection of the deterioration of patients in intensive and intermediate care units. *Journal of Computer Science and Technology*, 18(03):e25, Dec. 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, Marina Morán, Dolores Rexachs, and Emilio Luque. Factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. *XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, 2018.

[7] 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.

[8] 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.

[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] 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, Adolfy Hoisie, Dean Micron 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 Micron 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.

[11] Carolyn McGregor. Big data in neonatal intensive care. *Computer*, 46(6):54–59, 2013.

[12] 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.

Arquitecturas multiprocesador: Hardware, Software, Modelos, Métricas y Tendencias

De Giusti Armando[(1)(2)], Marcelo Naiouf[(1)], Tinetti Fernando[(1)(3)], Villagarcía Horacio[(1)(3)], Franco Chichizola[(1)], Laura De Giusti[(1)(3)], Enzo Rucci[(1)], Adrián Pousa[(1)], Victoria Sanz [(1)(3)], Montezanti Diego[(1)], Encinas Diego[(1)], Ismael Rodríguez[(1)], Sebastián Rodríguez Eguren[(1)], Erica Montes de Oca[(1)], Juan Manuel Paniego[(1)], Martín Pi Puig[(1)], César Estrebou[(1)], Leandro Libutti[(1)], Costanzo Manuel[(1)], Boggia Marcos[(1)],Javier Balladini[(4)]

[1]Instituto 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
[2]CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3]CICPBA – Comisión de Investigaciones Científicas de la Provincia de Buenos Aires
[4]Universidad Nacional del Comahue

{adegiusti,mnaiouf,fernando,hvw,francoch,ldgiusti,erucci,apousa,vsanz,dmontezanti,dencinas,ismael,seguren,emontesdeoca,jmpaniego,mpipuig,cesarest,llibutti, mcostanzo,mboggia}@lidi.info.unlp.edu.ar;
javier.balladini@gmail.com

Resumen

El eje de esta línea de I/D lo constituye el estudio de las arquitecturas multiprocesador que integran sistemas distribuidos y 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 computacional y energético.

Palabras clave: *Sistemas Paralelos. Clusters. Arquitecturas asimétricas. GPU, MIC, FPGA. Cloud Computing. Cloud robotics. Eficiencia energética. Resiliencia. E/S paralela.*

Contexto

Se presenta una línea de Investigación que 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 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 Ministerio de Educación, Cultura, Ciencia y Tecnología de la Nación.

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 [1].

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 [2].

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 [3]. 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*) [4][5][6].

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 [7].

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) [8][9]. 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 [10][11]. 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 [12].

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*" [13][14]. 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) [15].

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).

Recientemente, *Edge* y *Fog Computing* surgen como una evolución del modelo tradicional de Cloud en busca de disminuir el tiempo de respuesta de las soluciones y mejorar el aprovechamiento de recursos.

Resiliencia

En la actualidad, lograr sistemas resilientes resulta un verdadero desafío considerando el creciente número de componentes, la cercanía a los límites físicos en las tecnologías de fabricación y la complejidad incremental del software. La corrección de las aplicaciones y la eficiencia en su ejecución se torna más importante en HPC debido a los extensos tiempos de ejecución. En ese sentido, resulta interesante desarrollar estrategias de detección y recuperación de fallos, especialmente a través de librerías de software.

Entrada/Salida paralela

A pesar de los avances tecnológicos, las operaciones de E/S en los centros de supercómputo siguen siendo un cuello de botella para determinadas aplicaciones HPC. El rendimiento de un sistema dependerá de la carga de trabajo (patrones de E/S de las

aplicaciones) y de su configuración (hardware y software). Contar con herramientas que permitan modelar y predecir el comportamiento de este tipo de aplicaciones en HPC resulta fundamental para mejorar su rendimiento.

Dispositivos de bajo costo con capacidades para cómputo paralelo

En la actualidad se comercializan placas de bajo costo como Raspberry PI [16] u Odroid [17] 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. Técnicas de resiliencia.
- Exploración de nuevos lenguajes y modelos de programación para HPC.
- Cloud Computing para realizar HPC. Evaluación de perfomance en este tipo de arquitectura. Análisis del overhead por el software de administración del Cloud.
- Sistemas inteligentes de tiempo real distribuidos (Cloud Robotics). Edge y Fog Computing.
- 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. Modelado y estimación del consumo de potencia de arquitecturas HPC.

Resultados y Objetivos

Investigación experimental a realizar

- Desarrollar y evaluar algoritmos paralelos sobre nuevas arquitecturas. Analizar rendimiento, eficiencia energética y costo de programación.
- Analizar el overhead introducido por el sistema gestor del Cloud en un entorno de HPC para aplicaciones científicas de Big Data.
- Estudiar 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 [12][18].
- Optimizar 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.
- Desarrollar técnicas de tolerancia a fallas que permitan aumentar la resiliencia de sistemas paralelos y distribuidos.
- Desarrollar un modelo de E/S en HPC que permita predecir cómo los cambios realizados en los diferentes componentes del mismo afectan a la funcionalidad y al rendimiento del sistema.

Resultados obtenidos

- Se estudiaron técnicas de programación y optimización para la nueva generación de procesadores Xeon Phi (*Knights Landing*) [19].
- Se desarrollaron y analizaron soluciones híbridas CPU-GPU para el procesamiento de *pattern matching* [20].
- Se evaluó el uso de diferentes arquitecturas paralelas (CPU, Xeon Phi, GPU, FPGA) para distintos problemas de alta demanda computacional (procesamiento de grandes conjuntos de datos biológicos; cifrado de datos; simulación en astrofísica; entre otras)

considerando no sólo el rendimiento sino la eficiencia energética [21][22][23][24][25].
- Se estudió cómo paralelizar un modelo numérico sobre procesadores multicore [26].
- Se analizó la precisión en la predicción del consumo de energía por parte de la herramienta Intel RAPL en arquitecturas CPU [27].
- Se desarrolló y validó un modelo estadístico para consumo de potencia en placas RPi [28].
- Se desplegó un cluster de placas RPi y se estudió la viabilidad de su uso para cómputo de altas prestaciones considerando tanto pico de rendimiento como eficiencia energética [29].
- 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 [30].
- Se estudiaron los factores que afectan al consumo energético en las operaciones de checkpoint y restart en clusters [31].
- Se realizó un 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 [32][33].
- Se profundizó el desarrollo de un modelo de arquitectura para un sistema de virtualización distribuido [34].

Organización de Eventos

En el año 2018 se han organizado las VI Jornadas de Cloud Computing y Big Data (JCC&BD 2018) en Argentina, con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing. En junio de 2019 se organizarán las VII Jornadas de Cloud Computing y Big Data (JCC&BD 2019).

Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyeron: 1 tesis de maestría y 1 trabajo de Especialización. Al mismo tiempo se encuentran en curso 5 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

[1] Rucci, Enzo: “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
[2] Balladini, 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.
[3] 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).
[4] Lorencik D., Sincak P.: “Cloud robotics: Current trends and possible use as a service”. In: Applied Machine Intelligence and Informatics (SAMI), 2013 IEEE 11th International Symposium on , vol., no.,
[5] Guoqiang Hu, Wee Peng Tay, Yonggang Wen: “Cloud robotics: architecture, challenges and applications”. In: Network, IEEE, vol.26, no.3, pp.21-28. 2012.
[6] 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.
[7] Reinders, J., Jeffers, J., Sodani, A. "Intel Xeon Phi Processor High Performance Programming Knights Landing Edition". Morgan Kaufmann Publish-ers Inc., Boston, MA, USA, 2016
[8] 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
[9] Intel. "Intel Acquisition of Altera". Disponible en http://intelacquiresaltera.transactionannouncement.com
[10] Sean Settle: "High-performance Dynamic Programming on FPGAs with OpenCL". In: IEEE High Performance Extreme Computing Conference. 2013.
[11] Xilinx Inc. "SDAccel Development Environment". [Online]. Disponible en http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html
[12] 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.
[13] Shafer, J.: "I/O virtualization bottlenecks in cloud computing today". In: Proceedings of the 2nd conference on I/O virtualization (VIOV10). USA (2010).
[14] 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.
[15] Velte, A.T., Velte, T.J., Elsenpeter, R.: "Cloud Computing: A Practical Approach". McGraw Hill Professional. 2009.
[16] Raspberry PI. https://www.raspberrypi.org/
[17] Odroid http://www.hardkernel.com Accedido 21 de Marzo de 2016.
[18] 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.
[19] "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.
[20] V. Sanz, A. Pousa, M. Naiouf, and A. De Giusti, "Accelerating Pattern Matching with CPU-GPU Collaborative Computing", Proceedings of the International Conference on Algorithms and Architectures for Parallel Processing (ICA3pp 2018), ISBN: 978-3-030-05051-1, págs. 310-322, doi. https://doi.org/10.1007/978-3-030-05051-1_22, 2018.
[21] Rucci E., Garcia C., Botella G., De Giusti A., Naiouf M., Prieto-Matias M. (2018) SWIFOLD: Smith-Waterman Acceleration on Intel's FPGA with OpenCL for Long DNA Sequences. BMC Systems Biology2018 12(Suppl 5):96. DOI https://doi.org/10.1186/s12918-018-0614-6.
[22] SWIMM 2.0: enhanced Smith-Waterman on Intel's Multicore and Manycore architectures based on AVX-512 vector extensions Enzo Rucci, Carlos García, Guillermo Botella, Armando De Giusti, Marcelo Naiouf and Manuel Prieto-Matías. International Journal of Parallel Programming, Springer US, 2018. DOI: https://doi.org/10.1007/s10766-018-0585-7
[23] A. Pousa, V. Sanz, M. Naiouf, A. De Giusti. "Rendimiento del algoritmo AES sobre arquitecturas de memoria compartida", Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). *En prensa*, octubre de 2018.
[24] E. Montes de Oca, L. De Giusti, A. De Giusti, M. Naiouf. "Análisis de consumo energético en Cluster de GPU y MultiGPU en un problema de Alta Demanda Computacional.", Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). En prensa, octubre de 2018.
[25] M. Pi Puig, L. C. De Giusti, M. Naiouf, and A. E. De Giusti. "Are GPUs non-green devices?" Journal of computer science and technology (ISSN 1666-6038), vol. 18, num. 02, págs. 153-159, doi. 10.24215/16666038.18.e17, 2018.
[26] F. G. Tinetti, M. Pérez, A. Fraidenraich, and A. Altenberg, Proceedings of the International Conference on Parallel and Distributed Porcessing 2018 (PDPTA), ISBN: 978-1-60132-487-0, 2018.
[27] J. M. Paniego, S. Gallo, M. P. Puig, F. Chichizola, L. D. Giusti, and J. Balladini. "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.
[28] J. M. Paniego, L. Libutti, M. Pi Puig, F. Chichizola, L. De Giusti, M. Naiouf, A. De Giusti. "Modelado estadístico de potencia usando contadores de rendimiento sobre Raspberry Pi" . Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). *En prensa*, octubre de 2018.
[29] S. Rodriguez Eguren, F. Chichizola, E. Rucci. "Análisis del Uso de un Cluster de Raspberry Pi para Cómputo de Alto Rendimiento", Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). *En prensa*, octubre de 2018.
[30] 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.
[31] M. Morán, J. Balladini, D. Rexachs, E. Luque. "Factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters", Actas del XXIV Congreso Argentino de Ciencias de la

Computación (CACIC 2018). *En prensa*, octubre de 2018.

[32] M. Costanzo, M. Boggia, I. Rodríguez, and De Giusti Armando, "CLOUD ROBOTICs: Vehículo autónomo conectado a AWS" Actas de las VI Jornadas de Cloud Computing & Big Data (JCC&BD 2018), ISBN: 978-950-34-1659-4, págs. 14-22, 2018.

[33] M. Costanzo, M. Boggia, I. Rodríguez, A. De Giusti. "CLOUD ROBOTICs: Navegación de un vehículo autónomo en un entorno con obstáculos". Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). *En prensa*, octubre de 2018.

[34] P. Pessolani, F. G. Tinetti, T. Cortés, and S. Gonnet, Proceedings of the Ninth International Conference on Cloud Computing, GRIDs, and Virtualization (CLOUD COMPUTING 2018), págs. 1-11, 2018. pp.85-88. 2013. Accedido 21 de Marzo de 2016.

COMPUTACIÓN DE ALTO DESEMPEÑO: ALGORITMOS Y EVALUACION DE RENDIMIENTO EN DIFERENTES PLATAFORMAS

Marcelo Naiouf[(1)], Armando De Giusti[(1)(2)], Laura De Giusti[(1)(3)], Franco Chichizola[(1)], Victoria Sanz[(1)(3)], Adrián Pousa [(1)], Enzo Rucci[(1)], Silvana Gallo[(1)], Erica Montes de Oca[(1)], Emmanuel Frati[(1)], Mariano Sánchez [(1)], María José Basgall[(1)(2)], Adriana Gaudiani[(4)]

[1]Instituto 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

[2] CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas
[3]CIC – Comisión de Investigación Científica de la Provincia de Buenos Aires
[4] Universidad Nacional de General Sarmiento
{mnaiouf, degiusti, ldgiusti, francoch, vsanz, apousa, erucci, sgallo, emontesdeoca,fefrati, msanchez, mjbasgall}@lidi.info.unlp.edu.ar, agaudi@ungs.edu.ar

RESUMEN

El eje central de la línea de I/D 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 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; con el Departamento de Arquitectura de Computadores y Automática de la Universidad Complutense de Madrid; y con el grupo Soft Computing and Intelligent Information Systems (SCI2S) de la Universidad de Granada, 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, 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 [1].
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 [2] o los aceleradores basados en circuitos integrados reconfigurables (FPGAs, Field Programmable Gate Array) [3]. En la actualidad se comercializan placas de bajo costo como Raspberry PI [4] u Odroid [5] 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 [6], así como las características de scheduling en los mismos [7]. 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 [8]. 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 paralela 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 [9].

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 lenguajes y bibliotecas ya que aún no se cuenta con un standard, aunque puede mencionarse el uso de los tradicionales MPI, OpenMP y Pthreads [10][11] o los más recientemente explorados UPC, Chapel y Titanium del modelo PGAS [12].

La combinación de arquitecturas con múltiples núcleos con aceleradores dio lugar a plataformas híbridas con diferentes características. Más allá del acelerador utilizado, la programación de estas plataformas representa un desafío. Para lograr aplicaciones de alto rendimiento, los programadores enfrentan dificultades como: estudiar características específicas de cada arquitectura y aplicar técnicas de programación y optimización particulares de cada una, lograr un balance de carga adecuado entre los dispositivos de procesamiento y afrontar la ausencia de estándares 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 [13]. 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* [14]. 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 de la aplicación.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

- Investigar en temas de cómputo paralelo y distribuido de alto desempeño, en lo referido a los fundamentos y 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 diferentes 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.
▪ Respecto de las aplicaciones y temas estudiados, se trabajó fundamentalmente con los siguientes problemas:

➢ **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. En [15] se propuso un modelo híbrido para estructurar código a ser ejecutado sobre un sistema heterogéneo con múltiples cores y 1 GPU (utilizando todos los recursos disponibles). Utilizando este modelo se desarrolló un algoritmo paralelo de pattern matching para sistemas heterogéneos CPU-GPU. Los resultados revelaron que este algoritmo supera en rendimiento a trabajos previos, desarrollados para sistemas multicore y GPUs, para datos de tamaño considerable.

➢ **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 [16]. 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 [17]. 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 [18]. Por otra parte, la salida al mercado de las nuevas generaciones de procesadores Xeon (Skylake) y Xeon Phi (Knights Landing) brindó la oportunidad de evaluar su uso para búsquedas de similitud en bases de datos de proteínas [19][20]. A futuro, interesa explorar el uso de próximas generaciones de procesadores y aceleradores para esta aplicación.

➢ **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 [21]. 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 de enfermedades infecciosas transmitida por vectores utilizando modelos basados en agentes.** Se comenzó la adaptación de un modelo [22] de simulación sobre las enfermedades infecciosas dengue, zika y chikungunya, que permita abordar conocimientos que ayuden en la toma de decisiones y acciones de prevención. Dada la potencia de cómputo necesaria para su simulación se propone el uso de las plataformas cluster de multicore y GPU como infraestructura para disponer de simulaciones escalables y tiempo de simulación aceptables. Posteriormente, se llevará a cabo la validación del modelo desarrollado. Se propone realizar un estudio y análisis de performance en las arquitecturas propuestas. Además, se evaluar el consumo energético para determinar las relaciones entre escalabilidad, consumo y eficiencia.

➢ **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 [23]. También se han desarrollado diferentes alternativas de distribución de trabajado usando un Cluster de GPUs heterogéneas [24]. Además del uso de un Cluster de GPU se está utilizando MultiGPU, pero haciendo énfasis en el consumo energético. Se desarrolló un modelo de estimación de tiempo junto con un modelo de estimación de energía para el uso de una GPU, un Cluster de GPU y una MultiGPU. El trabajo experimental realizado se centró tanto en la performance como en el consumo energético de las distintas arquitecturas GPU empleadas, obteniéndose de este modo resultados que permitieron determinar las ventajas de una arquitectura sobre otra [25].

➢ **Problemas de optimización de simulación de sistemas dinámicos complejos mediante heurísticas**. La búsqueda de un conjunto de parámetros de entrada que optimicen el funcionamiento de un simulador de un sistema físico es un proceso de alto costo computacional que puede considerarse intratable y requiere de heurísticas que permitan disminuir el tiempo de ejecución. En los nuevos avances realizados en esta línea se aprovechó la continuidad en los valores de los parámetros físicos distribuidos sobre el dominio del sistema, de manera de realizar búsquedas de parámetros ajustados sobre espacios de búsquedas de tamaño mucho más reducidos que los utilizados en una primera etapa del trabajo. Esta metodología mucho más eficiente se puede extender a otros simuladores de fenómenos físicos y en particular con simuladores de inundaciones de ríos con las que se llevaron adelante las experiencias. Las experiencias se pueden correr de manera colaborativa beneficiándose ampliamente del uso de plataformas de clusters de procesadores [26] [27].

➢ **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 [28]. 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 [29]. Se ha ampliado para incorporar conceptos básicos de computación en la nube (Cloud Computing) [29].

➢ **HPC y Big Data.** Actualmente los escenarios Big Data son cada vez más comunes debido a la gran cantidad de datos que se genera constantemente. Para poder procesar dicha cantidad de datos y, en tiempos razonables de ejecución, es necesario utilizar cómputos de altas prestaciones. Además, actualmente existen diferentes herramientas pensadas para trabajar con este tipo de problemas. Una de ellas, y de las más ampliamente utilizadas, es conocida como Apache Spark. Este framework está basado en el paradigma MapReduce y, entre sus más importantes características se encuentran los mecanismos eficientes de tolerancia a fallos y el uso intensivo de memoria RAM que emplea. En esta línea de trabajo, se están adaptando técnicas

tradicionales aplicadas en pequeños datasets [30] como también desarrollando nuevas técnicas para el preprocesamiento de datos para problemas de clasificación no balanceada en Big Data, utilizando el framework Apache Spark.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática de la línea de I/D en el año 2018 se concluyó 1 Trabajo Final de Especialización y 1 Tesina de Grado de Licenciatura. Se encuentran en curso en el marco del proyecto 3 tesis doctorales, 2 de maestría, 3 trabajos de Especialización y 3 Tesinas de grado.
Se participa en el dictado de las carreras de Doctorado en Cs. Informáticas y Magíster 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 tesistas de diferentes Universidades realizan su trabajo con el equipo del proyecto.

5. BIBLIOGRAFÍA

[1]. Giles MB, Reguly I. “Trends in high-performance computing for engineering calculations”. Phil.Trans.R.Soc.A 372: 20130319. 2014. http://dx.doi.org/10.1098/rsta.2013.0319

[2]. Jeffers, James; Reinders, James. “Intel Xeon Phi Coprocessor High Performance Programming”. Morgan Kaufmann. 2013.

[3]. Sean Settle. “High-performance Dynamic Programming on FPGAs with OpenCL”. IEEE High Performance Extreme Computing Conf (HPEC ’13). 2013. https://doi.org/10.1016/j.physa.2017.04.159.

[4]. Raspberry PI. https://www.raspberrypi.org/ Accedido 21 de Marzo de 2016.

[5]. Odroid http://www.hardkernel.com Accedido 21 de Marzo de 2016.

[6]. 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.

[7]. 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.

[8]. McCool, Michael. “Structured Parallel Programming: Patterns for Efficient Computation”, Morgan Kaufmann, 2012

[9]. 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.

[10]. Chapman, B., Jost, G. & Van der Pas. “Using OpenMP – Portable Shared Memory Parallel Programming”. (2008). UK: MIT Press.

[11]. Hager, G. & Wellein, G. “Introduction to HPC for Scientists and Engineers”. (2011) EEUU: CRC Press.

[12]. De Wael, M; Marr, S; De Fraine, B; Van Cutsem, T; De Meuter, W. “Partitioned Global Addess Space Languages”. ACM Computing Surveys (CSUR) 47 (4), 2015.

[13]. Amazon Elastic Compute Cloud (Amazon EC2). http://aws.amazon.com/es/ec2/. Febrero 2013.

[14]. Balladini 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

[15]. V. Sanz, A. Pousa, M. Naiouf, and A. De Giusti, “Accelerating Pattern Matching with CPU-GPU Collaborative Computing”, Proceedings of the International Conference on Algorithms and Architectures for Parallel Processing (ICA3PP 2018), ISBN: 978-3-030-05051-1, págs. 310-322, doi. https://doi.org/10.1007/978-3-030-05051-1_22, 2018.

[16]. Enzo Rucci, Armando De Giusti, Marcelo Naiouf, Carlos García Sanchez, Guillermo Botella Juan, Manuel Prieto-Matías. “State-of-the-art in Smith-Waterman Protein Database Search”. 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.

[17]. 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.

[18]. E. Rucci, C. Garcia, G. Botella, A. De Giusti, M. Naiouf, and M. Prieto-Matias. “SWIFOLD:Smith-Waterman Implementation on FPGA with OpenCL for Long DNA Sequences”. BMC Systems Biology. ISSN: 1752-0509. In press. 2018.

[19]. 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”. 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.

[20]. Enzo Rucci, Carlos García, Guillermo Botella, Armando De Giusti, Marcelo Naiouf and Manuel Prieto-Matías. “SWIMM 2.0: enhanced Smith-Waterman on Intel's Multicore and Manycore architectures based on AVX-512 vector extensions”. International Journal of

Parallel Programming, Springer US, 2018. DOI: https://doi.org/10.1007/s10766-018-0585-7
[21]. E. Rucci, A. De Giusti, and M. Naiouf, “Blocked All-Pairs Shortest Paths Algorithm on Intel Xeon Phi KNL Processor: A Case Study”. 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.
[22]. Borges F., Gutierrez-Milla A., Suppi R., Luque E., de Brito Arduino M., “An agent-based model for assessment of aedes Aegypti pupal productivity”. Proceedings of the 2015 Winter Simulation Conference. ISBN: 978-1-4673-9743-8/15. IEEE. (2015)
[23]. 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.
[24]. 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.
[25]. Montes de Oca E. S. “Análisis de consumo energético en Cluster de GPU y MultiGPU en un problema de alta demanda computacional”. Trabajo Final Carrera de Especialista en Cómputo de Altas Prestaciones y Tecnología GRID. Facultad de Informática, Universidad Nacional de La Plata. (2018).
[26]. Trigila M., Gaudiani A., Luque E. (2018) “Agile Tuning Method in Successive Steps for a River Flow Simulator”. In: Shi Y. et al. (eds) Computational Science – ICCS 2018. ICCS 2018. Lecture Notes in Computer Science, vol 10862. Springer, Cham
[27]. Trigila, Mariano, Gaudiani, Adriana, Luque, Emilio. “Adjustment of a simulator of a complex dynamic system with emphasis on the reduction of computational resources”. Actas del Workshop. p. 142-147. 2018.
[28]. 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. 2018.
[29]. 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.
[30]. Basgall, M., Hasperué, W., Naiouf, M., Fernández, A., & Herrera, F. (2018). “SMOTE-BD: An Exact and Scalable Oversampling Method for Imbalanced Classification in Big Data”. Journal of Computer Science and Technology, 18(03), e23. https://doi.org/10.24215/16666038.18.e23.

IQR: Una medida estadística como modelo para la sintonización computacional

Caymes-Scutari Paola[1,2], Tardivo María Laura[1,3] Bianchini Germán[1], Méndez-Garabetti Miguel[1,]

[1]Laboratorio 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
[2]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)
[3]Departamento de Computación, Universidad Nacional de Río Cuarto, Río Cuarto, Córdoba, Argentina

pcaymesscutari@frm.utn.edu.ar, lauratardivo@dc.exa.unrc.edu.ar, gbianchini@frm.utn.edu.ar, mmendez@mendoza-conicet.gob.ar

RESUMEN

El Rango Intercuartil (o métrica IQR) es una medida estadística que cuantifica la dispersión de la muestra considerada, es decir, la variabilidad de la distribución de los elementos muestrales en base a algún parámetro de interés. En este proyecto, centrado en métodos que como muestra del espacio de búsqueda consideran poblaciones de individuos (muestra de elementos candidatos) se propone la utilización del rango intercuartil de los valores de aptitud de los individuos como un indicador de la tendencia a estancarse que tiene el algoritmo. El caso de estudio se refiere a la predicción de incendios forestales. Por lo tanto, la función de aptitud cuantifica el grado de coincidencia entre la predicción arrojada por cierto individuo y el incendio real. El IQR actuaría como un indicador de estancamiento y/o convergencia prematura y sería de utilidad para tomar decisiones sobre cuándo la predicción alcanzada es suficientemente buena, o bien ha alcanzado un cierto tope de calidad que no podrá mejorarse, y por consiguiente sea recomendable dar paso a una nueva población y una nueva generación. En resumen, se propone un modelo basado en el IQR para sintonizar de forma automática y dinámica el parámetro que regula la cantidad de generaciones del proceso evolutivo, a fin de evitar estancamiento y convergencia prematura.

Palabras clave: Rango Intercuartil, Evolución Diferencial, Algoritmos Evolutivos, Sintonización, Reducción de Incertidumbre.

CONTEXTO

El proyecto SIUTNME0004819 se encuentra en ejecución desde mayo de 2018, llevándose a cabo en el marco del LICPaD (Laboratorio de Investigación en Cómputo Paralelo/Distribuido) dentro del ámbito de la UTN-FRM. El mismo se encuentra enmarcado en el desarrollo de

un entorno denominado EDDSAP (Entorno de Especificación, Desarrollo y Sintonización de Aplicaciones Paralelas [1, 2]), 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 abordar la sintonización de los problemas de rendimiento que pueda presentar, mediante la consideración del rango intercuartil (IQR) como métrica de dispersión.

1. INTRODUCCIÓN

La predicción del comportamiento de un incendio forestal consiste en determinar cómo será la propagación del fuego sobre el terreno en un instante de tiempo futuro. Generalmente, los métodos de predicción implementan modelos que describen el comportamiento del fuego, y utilizan como dato de entrada un grupo de variables representando aquellos factores que condicionan la propagación. Entre ellas encontramos la velocidad y dirección del viento, la pendiente del terreno, el tipo de material combustible, la humedad de dicho material, etc. Lamentablemente, no es posible contar con los valores exactos para estos factores, debido a la imposibilidad de dotar todo el terreno forestal con instrumentos de medición y, sobre todo, debido a que algunos cambian dinámicamente durante el desarrollo del incendio [3]. Durante los últimos años se ha desarrollado una serie de métodos que pretenden reducir el impacto negativo que causa esta falta de conocimiento acerca del valor de las variables. Tal es el caso del método ESSIM-DE (Evolutionary Statistical Method with Island Model and Differential Evolution), el cual utiliza Estadística para obtener el patrón de comportamiento de la línea de fuego, el algoritmo evolutivo Evolución Diferencial [4], para orientar la búsqueda hacia mejores soluciones, y Cómputo Paralelo a través de un modelo de islas y múltiples poblaciones, con dos jerarquías de procesos que cooperan en la obtención de las predicciones, lo que permite explorar un amplio espacio de búsqueda y a la vez acelerar el proceso de predicción [5]. ESSIM-DE ha sido estudiado en el último tiempo con el fin de identificar aquellos aspectos que puedan ser factores limitantes de mejor rendimiento, en términos de calidad de las predicciones obtenidas y tiempo de respuesta. En este sentido, se propone dotar a ESSIM-DE con capacidades para su sintonización dinámica, lo que habilitará un mejor aprovechamiento de los recursos al adaptar durante la propia ejecución ciertos parámetros dependientes del problema o del estado del sistema. Dicho proceso de sintonización se logra a partir de cuatro fases que sucesivamente actúan sobre la aplicación [6]: *instrumentación* para anotar la aplicación con las métricas necesarias, *monitorización* para recolectar y clasificar las métricas, *análisis* para evaluar el estado del sistema en función de las métricas, y *sintonización* en sí misma para concretar los ajustes necesarios en la aplicación a fin de adaptar su comportamiento.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Esta temática se enmarca en la línea de investigación enfocada en el proceso de sintonización automática y dinámica de aplicaciones paralelas. El objetivo es detectar y modelar los problemas de rendimiento que pudiese traer

intrínsecamente aparejados el algoritmo de base considerado en la aplicación, a fin de darles tratamiento para paliar sus efectos negativos y obtener una ejecución más eficiente desde el punto de vista funcional, energético y de precisión en los resultados obtenidos. En este proyecto, se considera un método de predicción basado en evolución diferencial, y para abordar el proceso de sintonización hemos considerado dos posibles problemas asociados a su proceso evolutivo: la convergencia prematura y el estancamiento. La convergencia prematura es la situación en la cual la población converge a un óptimo local, debido a la pérdida de diversidad. Por su parte, el estancamiento es la situación en la cual el optimizador no es capaz de generar una nueva solución mejor que la anterior, aun cuando la población no hubiese convergido. Esto significa que, incluso cuando la población presente cierta diversidad, el optimizador es incapaz de encontrar mejores soluciones [7].

Así es que, para estos problemas, la sintonización se propone desde diversas aristas. Una de ellas consiste en la incorporación de un nuevo operador de reinicio (denominado **r**) de la población al inicio de cada paso de predicción, a fin de evitar el estancamiento global del proceso de búsqueda. Otra de las estrategias de sintonización propone la utilización de algún criterio para detectar, en cada generación, cuándo la búsqueda ha convergido lo suficiente, y por lo tanto poder interrumpir la ejecución del algoritmo a fin de obtener un resultado útil en un tiempo de ejecución menor. Para ello se propone analizar la tendencia decreciente de la distribución de la población a lo largo del proceso evolutivo, a fin de ajustar el rendimiento del método de acuerdo con el contexto de ejecución corriente, así como también al tipo de incendio considerado y a los diferentes parámetros que guían el proceso evolutivo en sí. Si bien en este proyecto el modelo de sintonización se aplica a un sistema de predicción de incendios, su utilización puede generalizarse para mejorar la performance de otras metaheurísticas, métodos de predicción y/o casos de estudio. El análisis a realizarse sobre las características de la población corriente ha de tener en cuenta fundamentalmente la diversidad poblacional, que puede modelarse y valorarse por medio de la métrica IQR. Por ello, se propone computar el Rango Intercuartil de los valores de aptitud de los individuos [8] a fin de cuantificar la dispersión de la población, es decir, la variabilidad de la distribución de la población en base a la aptitud de cada individuo, y así determinar la tendencia a la convergencia y/o el estancamiento.

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. Dicho objetivo será alcanzado a partir de otro aporte más particular, centrado en la definición de un modelo de rendimiento de ED basado en el rango intercuartil (IQR) que presenta la población. A su vez, 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, su participación en el proyecto, el estudio, experimentación y puesta en marcha de los procesos de análisis y sintonización de la evolución diferencial, será fundamental como elemento para su formación en el área de sintonización de aplicaciones paralelas, aplicable a su plan de trabajo. En el caso del Mg. Ing. Méndez Garabetti la temática también resulta transversal a su plan de tesis doctoral. Ambos estudiantes cursan el Doctorado en Ciencias de la Computación de la Universidad Nacional de San Luis. 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] 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.

[2] 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.

[3] 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

[4] Price K., Storn R., Lampinen J. (2005) Di□erential Evolution - A practical approach to global optimization. Springer-Verlag New York, Inc.

[5] Talbi, E. (2009) *Metaheuristics: From Design to Implementation*. John Wiley & Sons, Hoboken, New Jersey

[6] Naono, K., Teranishi, K., Cavazos, J. y Suda, R. (2010) *Software Automatic Tuning: From Concepts to State-of-the-Art Results*, Springer, New York.

[7] Lampinen, J. y Zelinka, I. (2000), "On the Stagnation of the Differential Evolution algorithm", *I.C. Soft Computing*, pp. 76-83.

[8] Healey, J. F. (2007) *The Essentials of Statistics: A Tool for Social Research*, Thomson/Wadsworth.

La Computación de Alta Performance como soporte a los sistemas altamente distribuidos

Diego Medel[1], María Murazzo[1], Ana Laura Molina[1], Federico Sánchez[1], Martín Cornejo[2], Nelson Rodríguez[1], Miguel Méndez-Garabetti[3], María Fabiana Piccoli[4]

[1] Departamento e Instituto de Informática - FCEFN - UNSJ

[2] Consejo Nacional de Investigaciones Científicas y Técnicas – CONICET

[3]Instituto de Investigaciones, Facultad de Informática y Diseño, Universidad Champagnat

[4]Universidad Nacional de San Luis

Complejo Islas Malvinas. Ignacio de la Roza y Meglioli. C.P. 5402. Rivadavia. San Juan, 0264 4234129

dmedel@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, amolina@unsj-cuim.edu.ar, fgsanchez@unsj-cuim.edu.ar, fmartin.cornejo@gmail.com, nelson@iinfo.unsj.edu.ar, mendez-garabettimiguel@uch.edu.ar, mpiccoli@unsl.edu.ar

RESUMEN

Debido a la aparición de nuevas tecnologías, dispositivos inteligentes, medios de comunicación y aplicaciones móviles, la cantidad de datos que se produce en la actualidad aumenta exponencialmente. Este aumento en la cantidad de datos demanda nuevas estrategias que permitan su almacenamiento, y análisis de manera eficiente; esto conlleva un cambio de paradigma en las arquitecturas de cómputo, los algoritmos y también los mecanismos de procesamiento.

Los sistemas de cómputo tradicionales han evolucionado a sistemas de alto rendimiento (HPC) para llevar a cabo cómputo intensivo y mejorar la velocidad de procesamiento.

La presente propuesta de investigación analizará y evaluará las ventajas de la computación de alto desempeño, aplicándolas a problemas que requieran un alto costo computacional en sistemas altamentes distribuidos, como son las aplicaciones de Internet de las Cosas.

Palabras claves: *High Computing Perfomance, IoT, Distribuited Systems*

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, aprobado 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 18 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 Facultad de Ciencias Exactas, Físicas y

Naturales de la Universidad Nacional de San Juan.

1. INTRODUCCION

Resolver problemas científicos, simular o modelar propuestas físicas o químicas han llevado a utilizar la computación como base a cálculos aritméticos o gráficos. Además, con la aparición de nuevas tecnologías, dispositivos inteligentes, aplicaciones móviles, redes sociales, la cantidad de datos que se produce en la actualidad aumenta exponencialmente. Se considera que el 90% de los datos existentes se han generado en los últimos dos años, esto está dando lugar a la Era del Exa y ZettaByte [1].

Debido a la necesidad de un alto costo de procesamiento y una gran cantidad de datos, los sistemas tradicionales, y en especial los de computo, han evolucionado a sistemas de cómputos de alto rendimiento (del *inglés High Performance Computing*, HPC). La evolución del procesamiento de información lleva varias décadas. Sin embargo, 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 lenguajes, librerías y algoritmos que las usen adecuadamente. Estos entornos son los ideales para *big data* debido a la creciente demanda y complejidad del cómputo necesario.

En [2], se describe los elementos de hardware que implican un sistema HPC físico. Desde el punto de vista de elementos de procesamiento, resalta procesadores *multicore* (8 núcleos o más), como así también, unidades de procesamiento gráfico (GPU). Siendo estas últimas protagonistas en aumentar la capacidad de computo de manera considerable. Además, distingue la interconexión de un clúster HPC según su escala o gama; los clústeres de gama baja o de menor escala tienden a utilizar variaciones de las redes Ethernet (Gigabit Ethernet), de gama media (media escala) redes Infiniband hasta 10 Gigabit Ethernet. Para los sistemas a gran escalar utilizan, por lo general, redes propietarias. Además, remarca el uso de módulos DDR3 DRAM de alta velocidad.

Tomando como arquitectura de base a un clúster, es posible combinar los paradigmas de memoria distribuida y memoria compartida, en un modelo de programación híbrida logrando aplicar una estrategia de paralización más efectiva mediante múltiples niveles de paralelismo y reducción del *overhead* de comunicación [3]. Librerías tales como MPI (*Message Passing Interface*) permiten el desarrollo de soluciones en el ámbito de memoria distribuida mediante la comunicación entre procesos por medio de funciones punto a punto o colectivas. OpenMP en ámbitos de memoria compartidas mediante la ejecución de *threads* y CUDA *(Compute Unified Device Architecture*) para arquitecturas paralelas como son las placas de videos mediante la ejecución concurrente de miles de *threads*. La programación hibrida consiste en desarrollar una solución en el que involucren ambos modelos de memoria mencionados.

Sin embargo, con el objetivo de mejorar los tiempos de procesamiento, se puede optar con la implementación de recursos de hardware con mayor potencia de cálculo, pero los costos de estos equipos son elevados, lo que dificulta su acceso a una gran cantidad de comunidades científicas. En [4], los autores definen a la nube (del inglés *cloud*) como un modelo de prestación de servicios, siendo la escalabilidad su principal característica. Es decir, los servicios son elásticos para el usuario, siendo transparente el funcionamiento del mismo. Es por ello, que la nube se ha convertido en un enorme repositorio de recursos computacionales, lo cual es una buena posibilidad para construir una plataforma para las aplicaciones que necesitan una gran cantidad de recursos. HPC se puede pensar como un servicio como muchos otros en la nube [5].

En la actualidad, la computación de alto desempeño en la nube es uno de los temas de investigación más actuales tanto por la comunidad científica como la empresarial. Los trabajos de HPC en la nube son una solución rentable que supera y reemplaza a los verdaderos clústeres dedicados. Grandes

empresas como Amazon Elastic Compute Cloud, Microsoft Azure and Google invierten millones de dólares en tecnología en la nube [6].

Desde hace unos años se empezó a hablar de Internet de las Cosas (del inglés Internet of Things, IoT) como un concepto de interconexión digital de objetos o dispositivos cotidianos. Cosas tales como vehículos, electrodomésticos o simplemente objetos (muebles, sensores, etc.) conectadas a la red. El concepto de IoT, tiene como objetivo hacer que internet sea aún más inmersiva y omnipresente. Además, permitir fácil acceso e interacción con una variedad amplia de dispositivos. IoT fomenta el desarrollo de una serie de aplicaciones que hacen uso de la enorme cantidad y variedad de datos generados por tales objetos para proporcionar nuevos servicios a ciudadanos, empresas y administraciones públicas. En [7], remarca que en los próximos 10 años la cantidad de dispositivos conectados crecerá en más de 50 mil millones. Es decir, la cantidad de datos generado por las aplicaciones de IoT seguirá creciendo y con ello la diversidad y la complejidad.

En [5], describen la necesidad de las aplicaciones de IoT en utilizar HPC. En su trabajo detalla ejemplos de aplicaciones relacionadas con el concepto de "mundo inteligente": *smart building management, smart logistics and smart manufacturing* son algunos de los ejemplos analizados. Los autores afirman la importancia en estudiar este vínculo HPC - IoT, ya que muchas aplicaciones de la IoT se caracterizan por la complejidad inherente al mundo físico, que conduce a problemas de HPC y, en particular, a problemas de optimización combinatoria.

2. LINEAS DE INVESTIGACION

Las tareas investigación se centrarán en el análisis de librerías, lenguajes y frameworks para el desarrollo de técnicas y/o algoritmos de cómputo intensivo sobre estructuras distribuidas, paralela e hibrida. Un análisis sobre aplicaciones que describan problemas científicos, simulaciones o aplicaciones de IoT que requieran un alto costo computacional. Estudiar los beneficios (desventajas) y características del uso de la Computación de Alto Desempeño en la nube, como así también de las aplicaciones que requieran gran cantidad de datos y costo computacional.

Profundizar el análisis de IoT y la necesidad de una computación de alta performance (HPC).

Además, debido a que los requerimientos de IoT en general son interactivos y los modelos de aplicaciones paralelas por lo general no satisfacen estos requerimientos es que se debe modelar el sistema de forma diferente. Por otro lado, los dispositivos IoT pueden fallar, la red presentar *delay* y otros problemas que hacen que los desarrollos de estos sistemas sean bastante diferentes de las aplicaciones científicas que se ejecutan en arquitecturas paralelas.

3. RESULTADOS OBTENIDOS

Resultados Obtenidos

Durante los últimos diez años se trabajó en el área de Computación de Alto Desempeño, en particular sobre análisis de diversas arquitecturas paralelas y distribuidas, tales como: Cloud Computing (públicos, híbridos y privados), clúster 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 trabajos de investigación en diferentes Congresos y Jornadas [3], [8]–[10], se realizaron tres publicaciones en revistas científicas y se transfirieron los resultados mediante conferencias en eventos científicos.

Algunos de estos trabajos [3], [8] comprendieron un análisis comparativo del desempeño de distintos modelos de programación paralela (memoria compartida, memoria distribuida e híbrido) en diversas configuraciones de los mismos. Los estudios estuvieron aplicados a problemas de álgebra lineal y a la simulación de propagación de incendios forestales. Los resultados obtenidos reflejan una mayor reducción del tiempo de

ejecución al utilizar un modelo de programación híbrida.

Se han aprobado diecisiete tesinas de grado y un trabajo de especialización y se incorporaron dos becarios de investigación categoría alumno y se encuentran en desarrollo tesinas de grado y de maestría. Además, en el marco de esta línea de investigación se han finalizado dos trabajos de tesina de grado [11][12].

Objetivos

El objetivo del grupo de investigación es analizar las ventajas de la computación de alto desempeño, analizando distintas arquitecturas y evaluando cómo se comporta en términos de rendimiento (performance) aplicaciones que requieran un costo computacional alto.

4. FORMACION DE RECURSOS HUMANOS

El equipo de trabajo de esta línea de investigación está formado por docentes-investigadores de las universidades de San Juan y Champagnat (Mendoza), un becario doctoral de CONICET y 2 alumnos de la UNSJ. Recientemente se han defendido dos tesinas de grado relacionadas al área [11] y [12]. Actualmente, se están realizando dos tesis doctorales relacionadas a esta línea de investigación.

Cabe destacar que el proyecto marco de la presente línea de investigación incluye investigadores de UNSL y UNLaR y además se estarían realizando dos maestrías y varias tesinas de grado.

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.

5. BIBLIOGRAFIA

[1] V. Acín *et al.*, *Architectures and methodologies for future deployment of multi-site Zettabyte-Exascale data handling platforms*, vol. 664. 2015.

[2] D. Chavarria-Miranda, Z. Huang, and Y. Chen, “High-performance computing (HPC): Application & use in the power grid,” *IEEE Power Energy Soc. Gen. Meet.*, pp. 1–7, 2012.

[3] M. Murazzo, N. Rodriguez, A. L. Molina, F. Sanchez, and D. Medel, “Análisis de desempeño de una arquitectura híbrida MPI/OpenMP mediante problemas de álgebra lineal,” in *IV Seminario Argentina - Brasil de Tecnologías de la Información y la Comunicación*, 2016, pp. 1–8.

[4] P. Mell and T. Grance, “The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology,” 2011.

[5] D. El Baz, “IoT and the need for high performance computing,” *Proc. - 2014 Int. Conf. Identification, Inf. Knowl. Internet Things, IIKI 2014*, no. Mic, pp. 1–6, 2014.

[6] R. Aljamal, A. El-Mousa, and F. Jubair, “A comparative review of high-performance computing major cloud service providers,” *2018 9th Int. Conf. Inf. Commun. Syst. ICICS 2018*, vol. 2018–Janua, pp. 181–186, 2018.

[7] J. Jadav, C. Tappert, M. Kollmer, A. M. Burke, and P. Dhiman, “Using text analysis on web filter data to explore K-12 student learning behavior,” *2016 IEEE 7th Annu. Ubiquitous Comput. Electron. Mob. Commun. Conf. UEMCON 2016*, pp. 1–5, 2016.

[8] A. L. Molina, F. Sanchez, N. Rodriguez, M. Murazzo, and M. Méndez-Garabetti, “Análisis de rendimiento de un simulador de

propagación de incendios forestales de ejecución paralela en diversos escenarios." Congreso Internacional de Ciencias de la Computación y Sistemas de Información, 2018.

[9] A. L. Molina Maturano, N. Rodriguez, M. Murazzo, F. Pinciroli, and M. Méndez-Garabetti, "FASES PARA EL DESARROLLO DE PROGRAMAS PARALELOS Y SU APLICACIÓN EN LA PREDICCIÓN DEL COMPORTAMIENTO DE INCENDIOS FORESTALES," in *XX Workshop de Investigadores en Ciencias de la Computación*, 2018, pp. 706–709.

[10] D. Medel, G. Fernández, N. Rodríguez, M. Murazzo, and S. Juan, "El modelo de Mobile Cloud Computing aplicado al desarrollo de un sistema de alertas."

[11] A. L. Molina Maturano, "Definición de fases para el desarrollo de programas paralelos pertenecientes al modelo de memoria compartida y su aplicación a la predicción del comportamiento de incendios forestales," Universidad Nacional de San Juan, 2018.

[12] F. Sanchez, "Paralelismo Híbrido para resolver problemas de cómputo intensivo," 2018.

Performance de arquitecturas multiprocesador: técnicas de modelado y simulación en HPC y Cloud Computing.

Diego Encinas[1,2], Jimena Jara[1], Daniel Rosatto[1], Román Bond[1], Martin Morales[1,3]

[1]Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

[2]Instituto 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, 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-BasedModeling and Simulation, ABMS). Cloud Computing. CloudSim.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación "Simulación y tecnología en Cómputo de Altas Prestaciones (High Performance Computing, HPC) para aplicaciones de interés social" de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 148/18. Además, el proyecto aporta al 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 y el Área de Computación del Instituto de Ciencias de la Universidad Nacional de General Sarmiento.

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 qué 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]. La idea general de la implementación consiste en la creación de una nueva SimEntity y sus agregaciones necesarias para manipular el espacio de memoria proporcionada por RamEntity, en tiempos de simulación. De tal forma que cada celda de RamEntity no sólo guarda datos, sino que también es posible que mantenga objetos interpretables como eventos para la nueva SimEntity que se denomina InitEntity.

Dichos eventos se planifican en tiempos de simulación, entonces se genera una nueva capa de ejecución en la simulación. La agrupación de eventos planificados para que sean ejecutados por InitEntity es definida como un ObjectCode, es decir lenguaje de máquina compilado listo para ser asignado a la pila de ejecución. En el caso del modelo de simulación estudiado, se establece que el código objeto se convierte en proceso. Ocurre cuando InitEntity lo recibe como dato del evento INIT_ADD_PR_ETC_RC del procesador de eventos de CloudSim, entonces es un proceso en estado listo. Cuando el procesador de eventos de InitEntity procesa un evento INIT_EXE_PR_ETC_RC puede tomar un proceso en estado listo, o también un proceso en estado bloqueado, si necesita un recurso no disponible y lo ejecuta. Los procesos que finalizan no se reasignan a la cola de procesos etc_rc. El código objeto es extensible para ser almacenado en el local storage para futuras entregas del modelo, es decir que durante la ejecución el proceso podría cambiar el código objeto almacenado en el local storage, en tiempos de simulación.
Lo destacable es que el entorno de la cola de procesos es en el espacio de memoria de RamEntity y que cada instrucción que compone a un código objeto del proceso es en realidad un evento con todos sus parámetros. La ejecución consiste en que el evento es programado en el motor de CloudSim, según el planificador del proceso Init. Cuando el proceso finaliza, el memory magnament de Init libera el fragmento de memoria asignado.

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, modelado e implementación de las operaciones típicas de E/S: read, write, open, close, flush.

- 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.

- Incorporación de tiempos de entrenamiento, obtenidos en AWS, en el simulador. Con esto se logra una salida más detallada y un método para validar tiempos y métricas del simulador con AWS.

- Implementación de comandos para ejecutar desde command center en NetLogo. Con esto se logró sintetizar el benchmark IOR correspondiente a la capa de aplicación de la pila de E/S, logrando introducir nuevos parámetros como tamaño de archivo y cantidad de nodos de E/S (metadata server y data server). De ejecutar este comando, se obtiene una nueva salida similar a la del benchmark IOR.

- Desarrollo de pruebas de conceptos con el modelo inicial del sistema de archivos paralelos utilizando el framework NetLogo.

- Los siguientes objetivos a tener en cuenta son: diferenciación de metadata server y data server, como dos parámetros diferentes; modelar e implementar la noción de IOPS, strip size y stride size.

- 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 InitEntity 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.

- Mejorar detalles de la simulación, cada SyntethicVm inicializará automáticamente el ObjectCode con el prefijo Init_ seguido del nombre de la SyntethicVm, utilizando las ventajas de reflection de java.

- Lograr crear una nueva capa de ejecución en tiempos de simulación con los recursos de las propias entidades.

- 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 2018 se han realizado 3 publicaciones nacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

En esta línea de I/D existe cooperación a nivel nacional e internacional. Hay un investigador realizando su Doctorado, 2 becarios de grado (EVC CIN) y 2

alumnos avanzados de grado colaborando en las tareas.

Referencias

1. Grama A, Gupta A, Karypis G, Kumar V. “Introduction to parallel computing”. SecondEdition. 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. Taufer 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 en IaaS privados y públicos.

Brian Galarza[1], Gonzalo Zaccardi[1,2], Román Bond[1], Federico Montes de Oca[1], Eduardo Maxit, Jorge Osio[1], David Duarte[1], Martin Morales[1,3], Diego Encinas[1,4]

[1]Instituto de Ingeniería y Agronomía - Universidad Nacional Arturo Jauretche

[2]Dirección Tecnológica – Universidad de Palermo

[3]Unidad CodApli - Facultad Regional La Plata - UTN

[4]Instituto de Investigación en Informática LIDI (III-LIDI) - Facultad de Informática - UNLP – Centro Asociado CIC

bgalarza@unaj.edu.ar, gzaccardi@unaj.edu.ar, roman.alejandro.b@gmail.com, federicomdo97@gmail.com, eduardomaxit5@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. Sistemas de Archivos en clústers.*

Contexto

Se presenta una línea de Investigación que es parte del Proyecto de Investigación "Simulación y tecnología en Cómputo de Altas Prestaciones (High Performance Computing, HPC) para aplicaciones de interés social" de la Universidad Nacional Arturo Jauretche (UNAJ), acreditado por resolución interna 148/18. Además, el proyecto aporta al Programa "Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social" de 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] como también OpenNebula [3].

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), autentificación (Keystone), Imagenes (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 [4] [5] [6].

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 [7].

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.

OpenNebula es un software de código abierto que permite el despliegue de IaaS. Busca reducir la complejidad generada por OpenStack y ofrece soporte con hypervisores tales como KVM y VMware vCenter.

OpenNebula clasifica a los nodos en dos tipos, Front – end los cuales entran en contacto con los usuarios y a su vez se comunican con los nodos de la infraestructura en los cuales se lanzaran las instancias y los nodos virtualizados los que a su vez deben contar con los paquetes correspondientes de storage, autenticación y networking para poder funcionar correctamente.

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.

Sistemas de Archivos Paralelos en clusters

Amazon Webservices [8], mediante el servicio EC2 (Amazon Elastic Compute Cloud) permite desplegar clústeres virtuales mediante instancias de VMs y almacenamiento para las mismas. Mediante este servicio y el correspondiente clúster conformado, se utiliza un sistema de archivos paralelo (PVFS2) que permite la gestión de datos particionados y distribuidos en los distintos nodos, mediante múltiples tareas de una aplicación ejecutada sobre el cluster.

PVFS2 utiliza una estructura de cliente-servidor. En función de su rol, existen tres tipos de nodos dentro de un cluster con este sistema de archivos: servidores de datos, servidores de metadatos y clientes, en donde cada uno de los nodos puede cumplir los tres roles.

PVFS2 [9] contiene una herramienta interna que permite volcar información en logs durante la ejecución de tareas dentro del sistema de archivos. Por cada uno de los nodos, el administrador del clúster puede obtener información de depuración (GOSSIP) conformado por registros de debug, de acceso, contadores de rendimiento y errores producidos durante la ejecución.

Asimismo, es posible especificar los datos a obtener en función de las distintas capas de PVFS2, del rol específico de cada nodo (cliente-servidor), así como también de otros factores relacionados con el funcionamiento interno y operaciones asociadas a la gestión de archivos.

Teniendo conocimiento acerca del funcionamiento del código fuente correspondiente a PVFS2 y del modo en el que realiza la escritura de registros en los logs, resulta factible realizar modificaciones que permitan obtener otro tipo de parámetros de interés a partir de esta herramienta.

Finalmente se propone el análisis, uso y configuración de distintas herramientas no invasivas para determinar la performance del sistema de archivos en clústers virtuales.

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.
- Sistemas de Archivos Paralelos.

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.

- Implementación de OpenNebula en un sistema con las mismas características que el implementado

por OpenStack con el fin de poder realizar pruebas en entornos similares.

- Utilización de Fuel para administrar OpenStack como sistema de administración de nube (Cloud Computing) a partir de la infraestructura de 2 nodos compute y e1 controller [10].

- Lograr escalabilidad agregando nuevos nodos compute a la infraestructura obteniendo un mayor performance en el sistema.

- Análisis 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 y OpenNebula para así comparar los resultados obtenidos de ambas infraestructuras y poder realizar un análisis del rendimiento en cada caso.

- Análisis y configuración de clusters virtuales.

- Análisis y configuración de herramientas no invasivas para la obtención de métricas en las distintas capas de software de los sistemas de archivos paralelos.

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 2018 se han realizado 3 publicaciones nacionales. Además, se encuentran en desarrollo y concluidas varias Prácticas Profesionales Supervisadas (PPS) con las que concluyen sus estudios los alumnos de Ingeniería en Informática.

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 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. OpenNebula. https://opennebula.org/. Febrero 2019
4. 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.
5. Zaccardi, G.; Galarza, B.; Encinas, D.;

Morales, M. “Implementación de Cloud Computing utilizando OpenStack”. 2° Congreso Nacional de Ingeniería Informática y Sistemas de Información (CoNaIISI 2014). Ciudad de San Luis, Argentina.

6. 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.
7. OpenStack Deployment Fuel. https://www.mirantis.com/products/mirantis-openstack-software/openstack-deployment-fuel/. Febrero 2016
8. Amazon Web Services (AWS)-Cloud Computing Services. https://aws.amazon.com Marzo 2019
9. T. PVFS2, “PVFS 2 File System Semantics Document,” tech. rep., PVFS Development Team, 2015
10. 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

Software y aplicaciones en Computación de Altas Prestaciones para el contexto de la UNdeC

Fernando Emmanuel FRATI, Jose TEXIER, Paula Cecilia RIVERA, Jonathan ALVAREZ, Fernanda CARMONA, Patricia FIGUEROLA, Francisco FRATI, Sebastián GUIDET, Roberto MILLON, Raul MORALEJO, Matías PEREZ, Emmanuel PORTUGAL, Donna RATTALINO, Alberto RIBA, Daniel ROBINS, Mara ROVERO, Javier RUITTI, Jorge TEJADA, Jusmeidy ZAMBRANO

Universidad Nacional de Chilecito
9 de julio 22, Chilecito, La Rioja, Argentina
{fefrati, jtexier, privera, jalvarez, fbcarmona, pfiguerola, ffrati, sguidet, rmillon, rmoralejo, mperez, eportugal, drattalino, ariba, drobins, mrovero, jruitti, jtejada, jzambrano}@undec.edu.ar

RESUMEN

En la actualidad, para abordar problemas de mayor tamaño y complejidad estudios de ciencia básica y aplicada utilizan Computación de Altas Prestaciones (HPC - High Performance Computing). El HPC permite mejorar la capacidad, velocidad y precisión en el procesamiento de datos. Con el proyecto que da origen a este trabajo se propone abordar seis estudios desde la perspectiva del HPC, para explorar los aspectos centrales del paralelismo aplicado desde las Ciencias de la Computación en otras disciplinas.

Algunas de los estudios que se abordan ya vienen desarrollándose en la Universidad Nacional de Chilecito, otros se inician a partir de cooperaciones con otras instituciones o para formalizar trabajos finales de postgrado. En todos los casos, el HPC será abordado a través de un proceso metodológico organizado para:

- Consolidar una infraestructura de experimentación, desarrollo y producción de soluciones a problemas de HPC
- Desarrollar las capacidades científico-tecnológicas del equipo
- Fomentar la vinculación y transferencia con los sectores académico, social y productivo

Cada problema abordado reúne entre sus integrantes, investigadores especialistas en la disciplina del estudio, investigadores de Ciencias de la Computación y estudiantes en sus últimos años de formación de grado. Se espera consolidar a corto plazo un grupo de investigación, desarrollo y transferencia que generará oportunidades de formación de recursos humanos, proveerá de servicios a la comunidad en el área de estudio y potenciará los vínculos de cooperación con otras instituciones.

Palabras clave: HPC, cómputo paralelo, aplicaciones, interdisciplinariedad, UNdeC.

CONTEXTO

La línea de investigación presentada es parte del proyecto "Software y aplicaciones en Computación de Altas Prestaciones" aprobado en el año 2018 por la Secretaría de Ciencia y Tecnología de la UNdeC para ser ejecutado desde marzo 2019 con una duración de 18 meses, convocatoria a proyectos de Investigación y Desarrollo. Además, en 2018 la UNdeC destinó fondos de PROMINF para la adquisición de equipamiento para el "laboratorio de sistemas paralelos". También el "Plan de mejoramiento de la función de I+D+i" (Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación) prevé financiamiento para la adquisición de equipamiento durante 2019 a fin de desarrollar las capacidades en HPC de la UNdeC. Estas iniciativas permitirán consolidar de una infraestructura de experimentación, desarrollo y producción de soluciones a problemas de HPC.

1. INTRODUCCIÓN

La informática tiene su origen en la necesidad de los distintos sectores de la sociedad de conseguir mayor velocidad, confiabilidad y precisión para resolver sus problemas. Sin embargo, la capacidad de solución a un problema dado encuentra su límite en los tiempos requeridos por sus algoritmos. Superar ese límite requiere que el problema sea abordado mediante cómputo paralelo. Normalmente, esto implica estudiar tres aspectos clave: hardware, aplicaciones y software.

Durante décadas, la industria respondió a la creciente demanda de mayor poder computacional incrementando exponencialmente el rendimiento de los procesadores. Esto fue posible gracias a tres tecnologías clave (escalado de transistores, diseño de la microarquitectura y jerarquías de memoria)[1]. Sin embargo, esta forma de obtener mayor poder de cómputo encontró barreras físicas (disipación de calor y eficiencia energía fundamentalmente), limitando el rendimiento de los microprocesadores y los sistemas en general [2]. Desde el año 2005, el escalado tecnológico se ha venido aprovechando para aumentar el número de cores dentro del chip, dando lugar a una importante variedad de arquitecturas (multicores, commodity clusters, GPGPU y Cloud), que permiten alcanzar enormes picos de rendimiento [3], [4].

No obstante, reducir los tiempos de procesamiento y obtener la mayor eficiencia de ese hardware requiere el diseño y desarrollo de algoritmos paralelos [5], [6]. Transformar un algoritmo secuencial en uno paralelo no es trivial. Un multicore está compuesto por dos o más unidades de procesamiento donde cada una puede ejecutar un flujo secuencial de instrucciones. Si estos flujos secuenciales de instrucciones corresponden a tareas de un mismo programa concurrente, entonces se trata de una ejecución paralela. En general estos procesos necesitan algún mecanismo para comunicar resultados parciales entre sí. Dependiendo de la arquitectura de cómputo, se consigue a través del uso de variables compartidas o del paso de mensajes entre procesos. Una transformación 'implícita' o transparente es deseable, pero el costo es una pérdida importante de rendimiento [7]. Por otro lado, debido a que las arquitecturas de cómputo cambian constantemente y a fin de reducir el esfuerzo necesario para migrar de una plataforma a otra, es deseable poder expresar los programas de manera independiente a la máquina. Por ello, el programador recurre a librerías estándares para expresar explícitamente el paralelismo: OpenMP, Pthreads, CUDA, OpenCL, MPI [8].

La secuencia cronológica en que se ejecutan las comunicaciones entre los procesos se conoce como historia del programa. Las suposiciones de orden de ejecución entre instrucciones heredadas del modelo de programación secuencial ya no son válidas, obligando al programador a utilizar algún mecanismo de sincronización para garantizar estados consistentes del programa. En este contexto, la correctitud de los algoritmos es más difícil de garantizar que en la computación serial. Frecuentemente el programador se equivoca al sincronizar los procesos, dando lugar a nuevos errores de programación: deadlocks, condiciones de carrera, violaciones de orden, violaciones de atomicidad simple y violaciones de atomicidad multivariable. Los métodos tradicionales de depuración de programas y las técnicas de sintonización de rendimiento existentes en la programación secuencial no se trasladan directamente a la programación paralela, requiriendo el uso de herramientas específicas [9]. Al diseñar una solución serial, existe un modelo teórico que permite estimar el desempeño de los programas antes de escribirlos. La evaluación del sistema paralelo (entendiendo por sistema al conjunto de software y hardware) se realiza a través de distintas métricas: tiempo de ejecución, speedup, eficiencia y overhead [10].

Las soluciones paralelas están tan estrechamente vinculadas con el hardware subyacente que dificultan enormemente

conseguir portabilidad de rendimiento: una solución pensada para una máquina paralela en memoria distribuida tendrá diferente eficiencia en una de memoria compartida, o incluso en otra máquina de memoria distribuida [11], [12].

Existen muchos aspectos que requieren ser tomados en cuenta al diseñar la solución paralela: tamaño del problema, división de datos o tareas, balance de carga, requerimientos de memoria, precisión de los cálculos, comunicaciones y sincronización entre procesos, errores de concurrencia, detección y tolerancia a fallos entre los más relevantes. La complejidad de los problemas requiere habilidades especiales de los desarrolladores: dominio de múltiples paradigmas de programación y frecuentemente múltiples lenguajes, conocimientos de redes y comprensión de la concurrencia y sus consecuencias. Por todo esto, se considera de gran interés el estudio de estos temas para el desarrollo de capacidades científico- tecnológicas en la UNdeC que favorezcan el trabajo interdisciplinario en la institución.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se propone abordar las siguientes líneas de I+D desde la perspectiva del HPC como eje central:

- Análisis de la diversidad molecular de microorganismos del suelo [13]–[18]. Estudio e implementación de algoritmos que contribuyan a reducir los tiempos de procesamiento y aumentar la capacidad de análisis referidos a este campo de la bioinformática, a fin de profundizar en el estudio de la diversidad molecular de microorganismos del suelo asociados a cultivos regionales.
- Detección y depuración de errores de concurrencia en programas paralelos [9], [19]–[22]. Estudio de técnicas y estrategias para implementar mecanismos de detección de errores de concurrencia robustos y confiables que contribuyan al diseño de programas paralelos libres de errores.
- Identificación biométrica masiva mediante venas del dedo usando redes de aprendizaje extremo (ELM) [23]–[27]. Estudio de técnicas de computación paralela para mejorar la eficiencia de identificación biométrica masiva basada en venas de dedo, para la aceleración del preprocesamiento y extracción de características biométricas, y el diseño de algoritmos de ELM mejorados que manejen eficientemente lotes de datos de gran tamaño.
- Servicios basados en lingüística computacional para análisis de texto [28]–[34]. Estudio sobre modelos computacionales que reproduzcan aspectos del lenguaje humano, con el fin de realizar análisis lingüísticos como servicios para el Centro de Escritura en la UNdeC.
- Documentos inteligentes a través del Blockchain [35]–[38]. Estudio de la tecnología blockchain para garantizar la integridad de documentos universitarios.
- Nodo de información meteorológica [39], [40]. Estudio, diseño e implementación de algoritmos para reducir los tiempos de procesamiento, aumentar la capacidad de análisis y favorecer la escalabilidad de aplicaciones de análisis y proyección de datos climáticos.

3. RESULTADOS ESPERADOS

Se consolidará un grupo de investigación, desarrollo y transferencia a la comunidad con capacidad para abordar problemas de computación de altas prestaciones, que generará oportunidades de formación de recursos humanos, proveerá de servicios a la comunidad en el área de estudio y potenciará los vínculos de cooperación con otras instituciones. Con respecto a las líneas propuestas, se espera:

- Establecer las oportunidades de mejora a través del paralelismo de desempeño, tiempo de respuesta o capacidad de procesamiento de cada subproyecto abordado.

- Determinar la configuración de recursos de cómputo que mejor se ajuste a cada subproyecto.
- Conocer las propuestas de la comunidad académica y científica para resolver el problema.
- Definir los criterios de diseño de software que conduzcan a una mejora de la aplicación.
- Obtener productos de software que puedan ser ejecutados en la plataforma de HPC como servicios para las áreas disciplinares vinculadas.
- Mejorar la visibilidad del grupo de I+D a nivel regional, nacional e internacional.
- Aumentar la cantidad de trabajos de finalización de grado, tesis de postgrado y actividades específicas en temas de HPC en la UNdeC.
- Determinar el éxito del proceso de I+D.

4. FORMACIÓN DE RECURSOS HUMANOS

Cinco miembros del equipo poseen formación de postgrado a nivel de doctorado, uno de ellos es especialista en Cómputo de Altas Prestaciones. Cinco miembros se encuentran en su etapa final para obtener el grado de maestría en Informática, tres de los cuales desarrollan como tesis temas abordados por esta propuesta. Dos de estas tesis de maestría están siendo codirigidas por docentes de la Universidad Católica de Maule (Chile). Cada línea I+D propuesta integra al menos un docente investigador experto en el campo de cada estudio específico. Todos los temas propuestos se trabajan con estudiantes de grado de las carreras Ingeniería en Sistemas y Licenciatura en Sistemas de la UNdeC (ambas acreditadas por CONEAU). Los docentes forman parte de los equipos de diversas asignaturas de estas carreras, entre las que se encuentran programación, arquitecturas de computadoras y arquitecturas paralelas. Nueve docentes se encuentran categorizados en el programa de incentivos.

5. BIBLIOGRAFÍA

[1] S. Borkar y A. A. Chien, «The Future of Microprocessors», *Commun ACM*, vol. 54, n.º 5, pp. 67–77, may 2011.

[2] K. Ahmed y K. Schuegraf, «Transistor wars», *IEEE Spectr.*, vol. 48, n.º 11, pp. 50-66, nov. 2011.

[3] V. V. Kindratenko *et al.*, «GPU clusters for high-performance computing», en *2009 IEEE International Conference on Cluster Computing and Workshops*, 2009, pp. 1–8.

[4] J. Jeffers, J. Reinders, y A. Sodani, *Intel Xeon Phi Processor High Performance Programming: Knights Landing Edition*. Morgan Kaufmann, 2016.

[5] G. R. Andrews, *Foundations of Multithreaded, Parallel, and Distributed Programming*. Addison Wesley, 2000.

[6] A. Grama, A. Gupta, G. Karypis, y V. Kumar, *Introduction to Parallel Computing - Second Edition*. Pearson Education and Addison Wesley, 2003.

[7] G. Hager y G. Wellein, *Introduction to high performance computing for scientists and engineers*. Boca Raton, FL: CRC Press, 2011.

[8] J. Dongarra *et al.*, *Sourcebook of parallel computing*, vol. 3003. Morgan Kaufmann Publishers San Francisco, 2003.

[9] F. E. Frati, «Software para arquitecturas basadas en procesadores de múltiples núcleos», Tesis, Facultad de Informática, 2015.

[10] X.-H. Sun, «Scalability versus execution time in scalable systems», *J. Parallel Distrib. Comput.*, vol. 62, n.º 2, pp. 173–192, 2002.

[11] C. Leopold, *Parallel and Distributed Computing: A survey of Models, Paradigms and approaches*. John Wiley & Sons, Inc., 2001.

[12] S. Ghosh, *Distributed Systems: An Algorithmic Approach*, 1.ª ed. University of Iowa, Iowa City, USA: Chapman and Hall/CRC, 2006.

[13] M. S. De, M. Prager, R. E. Naranjo, y O. E. Sanclemente, «El suelo, su metabolismo, ciclaje de nutrientes y prácticas agroecológicas», *Agroecología*, vol. 7, n.º 1, pp. 19–34, 2012.

[14] R. F. Cerrato y A. Alarcón, «La microbiología del suelo en la agricultura sostenible», *Cienc. -Sum*, vol. 8, n.º 2, pp. 175–183, 2001.

[15] J. P. Hulsenbeck y F. Ronquist, «MrBayes: Bayesian inference of phylogeny», *Bioinformatics*, vol. 17, pp. 754–755, 2001.

[16] A. Stamatakis, «RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies», *Bioinformatics*, vol. 30, n.º 9, pp. 1312-1313, may 2014.

[17] M. A. Suchard, P. Lemey, G. Baele, D. L. Ayres, A. J. Drummond, y A. Rambaut,

«Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10», *Virus Evol.*, vol. 4, n.º 1, ene. 2018.

[18] D. L. Swofford, «PAUP*: phylogenetic analysis using parsimony (*and Other Methods), version 4», *Sinauer Assoc. Sunderland Mass.*, 2003.

[19] S. V. Adve y H.-J. Boehm, «Memory models: a case for rethinking parallel languages and hardware», *Commun. ACM*, vol. 53, n.º 8, pp. 90–101, 2010.

[20] M. Ronsse y K. De Bosschere, «RecPlay: a fully integrated practical record/replay system», *ACM Trans Comput Syst*, vol. 17, n.º 2, pp. 133–152, 1999.

[21] S. Lu, S. Park, E. Seo, y Y. Zhou, «Learning from mistakes: a comprehensive study on real world concurrency bug characteristics», *SIGARCH Comput Arch. News*, vol. 36, n.º 1, pp. 329–339, 2008.

[22] S. Lu, J. Tucek, F. Qin, y Y. Zhou, «AVIO: detecting atomicity violations via access interleaving invariants», *SIGPLAN Not*, vol. 41, n.º 11, pp. 37–48, 2006.

[23] K. Wang, H. Ma, O. P. Popoola, y J. Liu, «Finger vein recognition», en *Biometrics*, InTech, 2011.

[24] D. Ezhilmaran y P. R. B. Joseph, «A STUDY OF FEATURE EXTRACTION TECHNIQUES AND IMAGE ENHANCEMENT ALGORITHMS FOR FINGER VEIN RECOGNITION», p. 8, 2015.

[25] G.-B. Huang, Q.-Y. Zhu, y C.-K. Siew, «Extreme learning machine: Theory and applications», *Neurocomputing*, vol. 70, n.º 1-3, pp. 489-501, dic. 2006.

[26] G. Huang, G.-B. Huang, S. Song, y K. You, «Trends in extreme learning machines: A review», *Neural Netw.*, vol. 61, pp. 32-48, ene. 2015.

[27] A. Akusok, K. Björk, Y. Miche, y A. Lendasse, «High-Performance Extreme Learning Machines: A Complete Toolbox for Big Data Applications», *IEEE Access*, vol. 3, pp. 1011-1025, 2015.

[28] AMPLN, «Asociación Mexicana para el Procesamiento del Lenguaje Natural Main/Home Page». [En línea]. Disponible en: https://www.ampln.org/. [Accedido: 31-ago-2018].

[29] M. Vallez y R. Pedraza, «El Procesamiento del Lenguaje Natural en la Recuperación de Información Textual y áreas afines», *Hipertext Net*, 2007.

[30] J. Texier, F. E. Frati, F. B. Carmona, A. E. Riba, M. Pérez, y J. Zambrano, «La gestión de la información en abierto, vehículo importante para maximizar la visibilidad web», presentado en XVIII Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina), 2016.

[31] P. Gamallo Otero, J. C. Pichel Campos, M. García González, J. M. Abuín Mosquera, y T. Fernández Pena, «Análisis morfosintáctico y clasificación de entidades nombradas en un entorno Big Data», 2014.

[32] S. W. D. Chien, C. P. Sishtla, S. Markidis, J. Zhang, I. B. Peng, y E. Laure, «An Evaluation of the TensorFlow Programming Model for Solving Traditional HPC Problems», en *International Conference on Exascale Applications and Software*, 2018, p. 34.

[33] H. Guan, X. Shen, y H. Krim, «Egeria: A Framework for Automatic Synthesis of HPC Advising Tools Through Multi-layered Natural Language Processing», en *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, New York, NY, USA, 2017, pp. 10:1–10:14.

[34] P. Suber, *Ensuring open access for publicly funded research*. British Medical Journal Publishing Group, 2012.

[35] M. U. Wasim, A. A. Ibrahim, P. Bouvry, y T. Limba, «Law as a service (LaaS): Enabling legal protection over a blockchain network», en *Smart Cities: Improving Quality of Life Using ICT & IoT (HONET-ICT), 2017 14th International Conference on*, 2017, pp. 110–114.

[36] A. Preukschat, *Blockchain: la revolución industrial de internet*. Gestión 2000, 2017.

[37] L.-Y. Yeh, P. J. Lu, y J.-W. Hu, «NCHC blockchain construction platform (NBCP): rapidly constructing blockchain nodes around Taiwan», en *Digital Libraries (JCDL), 2017 ACM/IEEE Joint Conference on*, 2017, pp. 1–2.

[38] H. Dai *et al.*, «TrialChain: A Blockchain-Based Platform to Validate Data Integrity in Large, Biomedical Research Studies», *ArXiv Prepr. ArXiv180703662*, 2018.

[39] A. Botta, W. De Donato, V. Persico, y A. Pescapé, «On the integration of cloud computing and internet of things», en *Future internet of things and cloud (FiCloud), 2014 international conference on*, 2014, pp. 23–30.

[40] P. Yue, H. Zhou, J. Gong, y L. Hu, «Geoprocessing in cloud computing platforms–a comparative analysis», *Int. J. Digit. Earth*, vol. 6, n.º 4, pp. 404–425, 2013.

Seguridad

Informática

Análisis de la seguridad de los datos en Internet de las Cosas usando tecnología Blockchain

Mg. Jorge Eterovic; Esp. Marcelo Cipriano; Lic. Luis Torres

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; torreslu@ar.ibm.com

RESUMEN.

El mayor reto al que se enfrenta la seguridad de IoT (Internet of Things, Internet de las Cosas) procede de la propia arquitectura del ecosistema actual, que se basa por completo en un modelo centralizado conocido como cliente/servidor. Todos los dispositivos se identifican, autentican y conectan a través de servidores en la nube.

La conexión entre los dispositivos tiene que realizarse a través de la nube, aunque se encuentren separados por tan solo unos pocos metros. Aunque este modelo ha interconectado dispositivos informáticos durante décadas, no podrá responder a las crecientes necesidades de los enormes ecosistemas de IoT del futuro.

La seguridad de este modelo se basa en la existencia de terceras entidades de confianza que emiten certificados digitales a un determinado costo.

El Blockchain es una de las tecnologías más innovadoras de nuestro tiempo y su uso viene ganando interés desde su aparición, gracias a su capacidad para asegurar la integridad de las transacciones y la autenticidad entre cualquier entidad conectada a Internet, de manera descentralizada, lo que significa que no hay un servidor maestro que albergue toda la cadena de transacciones. En su lugar, los nodos participantes tienen una copia de la cadena.

La gran ventaja de Blockchain es que es público. Todos los que participan pueden ver los bloques y las transacciones almacenadas en ellos. La aplicación de esta tecnología en el campo de IoT busca garantizar la seguridad y la privacidad de los datos en la interconexión digital de dispositivos físicos a través de Internet, ya que la cadena de bloques (Blockchain) es descentralizada, por lo que no hay una única autoridad que apruebe las transacciones o defina reglas específicas para la aceptación de transacciones.

Palabras Clave:
Seguridad en Internet de las Cosas. Blockchain. Convergencia IoT-Blockchain.

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 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 (2019-2020).

1. INTRODUCCIÓN.

Durante la última década, el IoT (Internet of Things) se ha ido introduciendo gradualmente en nuestras vidas gracias a la disponibilidad de sistemas de comunicación inalámbricos [1].

El paradigma de IoT abarca muchos conceptos; dispositivos inteligentes que recopilan datos del entorno, muchas tecnologías diferentes para permitir su conexión, servicios y estándares, y todos los elementos que participan [2], [3].

IoT puede traer muchos beneficios a la sociedad de muchas maneras diferentes, pero es muy importante estar atento para que se encuentre la mejor solución para proteger la privacidad de los datos [4] [5].

El gran desafío del IoT es encontrar un entorno de comunicación confiable que garantice la seguridad de los datos transmitidos entre todos los dispositivos conectados [6].

Una de las posibles soluciones podría ser la convergencia entre la tecnología IoT y Blockchain. Esta hipótesis se analizará en este proyecto de investigación [7].

Hay tres elementos principales a los que se hace referencia en una arquitectura de IoT [8] [10]:

1. Cosas: Dispositivos que tienen un medio de conectarse a una red más amplia.
2. Red: Conecta los múltiples dispositivos a la nube.
3. Nube: Los servidores remotos en un centro de datos cuya función es consolidar y almacenar los datos de forma segura.

La tecnología Blockchain tiene el potencial de asegurar este entorno, gracias a la criptografía, que todos los usuarios tengan una clave secreta y única a través de la cual acceden a sus datos [9].

Los protocolos abiertos basados en Blockchain pueden estandarizar el uso de los protocolos de comunicaciones y geolocalización de todos los tipos de sensores, tan comunes por ejemplo, en las líneas de montaje del sector industrial para automatizar los datos de forma segura (Industria 4.0) [11].

Muchos sectores se beneficiarán de esta tecnología, como el del cuidado de la salud o los sectores de transporte y logística, que también podrán obtener grandes beneficios de su uso para integrar las ventajas del IoT [8].

Las ventajas que ofrece incluyen la descentralización y la transparencia de la información, ya que todos los actores están en el mismo nivel jerárquico, evitando que un organizador principal use incorrectamente los datos [3].

Aunque actualmente no se puede garantizar el entorno deseado de una manera completamente segura, muchas empresas y organizaciones trabajan para mejorar cada vez más esta cobertura de Internet de las cosas. No es imposible evitar los ataques, pero los hace mucho más difíciles de producir.

Muchos estudios se centran actualmente en el uso de cifrado homomórfico. En el pasado, ya se ha demostrado que este cifrado era factible, pero requiere mucho más tiempo que los procesos convencionales, y hay que seguir investigando para mejorarlo.

Por el momento no existe un desarrollo sólido para mantener la seguridad deseada en este entorno y los ataques a los dispositivos son el principal desafío que tenemos por delante [4].

Este proyecto de investigación se centra en la necesidad de revisar y estudiar toda la información a través de diversas investigaciones llevadas a cabo en el campo de IoT, Blockchain y su convergencia, y analizar cuáles son las mejores soluciones para minimizar las amenazas y vulnerabilidades de las principales tecnologías que forman parte de Internet de cosas.

El resultado esperado es reunir la información clave del paradigma de IoT relacionado con la seguridad y los mejores mecanismos para garantizar la integridad y la privacidad de los datos en IoT.

Otro objetivo es permitir que otras personas interesadas en este campo accedan fácilmente a la información y el conocimiento recogidos en este proyecto.

Finalmente, se propondrán algunos escenarios donde podría converger la tecnología IoT y Blockchain en el futuro.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO.

La metodología para llevar a cabo este proyecto es buscar información en artículos de investigación y publicaciones para estudiar las amenazas y vulnerabilidades más importantes que se han presentado hasta el momento en el paradigma de IoT.

También se estudiará como aporta a la seguridad la tecnología Blockchain.

Luego, se analizará la convergencia de ambas tecnologías y como se puede resolver el problema de garantizar la privacidad y seguridad de los datos en IoT usando Blockchain para generar entornos de confianza que pueden prevenir futuros ataques.

Finalmente se redactará un informe con los resultados obtenidos.

3. RESULTADOS OBTENIDOS/ ESPERADOS.

Los objetivos de este proyecto de investigación son:

- Exponer en detalle los conceptos de IoT (Internet de las cosas) y Blockchain y hacer un estudio de su impacto en la sociedad y sus perspectivas de futuro como posibles aplicaciones.
- Hacer un estudio de los artículos de investigación y publicaciones que relacionen estos conceptos, destacando las contribuciones más importantes.
- Analizar la privacidad en el intercambio de datos en las principales tecnologías de IoT y cuáles son las mejores contramedidas para evitar las posibles amenazas.
- Analizar los desafíos y oportunidades de la convergencia entre IoT y la tecnología Blockchain.
- Proponer la convergencia de las tecnologías IoT y Blockchain para aplicaciones seguras.

4. FORMACIÓN DE RECURSOS HUMANOS.

El equipo de investigadores pertenece al cuerpo docente de Tecnologías Aplicadas de la Facultad de Ingeniería, específicamente al área de la Seguridad Informática, de la Universidad del Salvador.

A este proyecto, que recién inicia, se incorporarán próximamente un docente investigador con amplia experiencia en la industria y 2 alumnos que se encuentran promediando la carrera de Ingeniería en Informática.

Esto redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes investigadores, como así también sembrará las bases para la investigación a futuro, a través de la participación de alumnos de la Facultad de Ingeniería.

5. BIBLIOGRAFÍA.

[1] F. Mattern and C. Floerkemeier; "From the Internet of Computers to the Internet of Things"; ACM Digital Library; From Active Data Management to Event-Based Systems and More; 2010. http://www.vs.inf.ethz.ch /publ/papers/Internet-of-things.pdf.

[2] R. Minerva, A Biru, D. Rotondi; "Towards a definition of the Internet of Things (IoT)"; IEEE Xplore Digital Library; 2015. https://iot.ieee.org/images/files/pdf/

IEEE_IoT_Towards_Definition_Internet_of_ Things_ Revision1_27MAY15.pdf

[3] Shaddad Abdul-Qawy, P. P. J, E. Magesh, T. Srinivasulu; "The Internet of Things (IoT): An Overview"; Directory of Open Access Journals; International Journal of Engineering Research and Applications; V.5, N. 12; 2015.

[4] D.Mendez, I. Papapanagiotou, B Yang; "Internet of Things: Survey on Security and Privacy"; Cornell University Library; 2017. https://arxiv.org/abs/1707.01879

[5] J. Granjal, E. Monteiro, J. Sá Silva; "Security for the Internet of Things: A Survey of Existing Protocols and Open Research Issues"; IEEE Xplore Digital Library; IEEE Communications Surveys & Tutorials, 2015, Volume 17, Number 3; http://0-ieeexplore.ieee.org.cataleg.uoc.edu/ document/7005393

[6] A. Bahga, V. K. Madisetti, "Blockchain Platform for Industrial Internet of Things"; Scientific Research; Vol.9, No.10, October 2016. https://www.scirp.org/Journal/Paper Information.aspx?PaperID=71596

[7] E. Gaetani, L. Aniello, R. Baldoni, F. Lombardi, A Margheri, and V. Sassone; "Blockchain-based Database to Ensure Data Integrity in Cloud Computing Environments"; First Italian Conference on Cybersecurity (ITASEC17); Venice, Italy; 2017. http://ceur-ws.org/Vol-1816/paper-15.pdf

[8] G.Zyskind, O. Nathan, A. Sandy Pentland; "Decentralizing Privacy: Using Blockchain to Protect Personal Data"; IEEE Xplore Digital Library; Security and Privacy Workshops (SPW); 2015. http://ieeexplore.ieee.org/ document/ 7163223/

[9] T. Tuan Anh Dinh, R. Liu; "Untangling Blockchain: A Data Processing View of Blockchain Systems"; Cornell University Library; 2017. http://www.comp.nus.edu.sg/ ~ooibc/blockchainsurvey.pdf

[10] K. Christidis, "Blockchains and Smart Contracts for the Internet of Things" IEEE Xplore Digital Library, IEEE Access, Volume 4; 2016. http://0-ieeexplore.ieee.org.cataleg. uoc.edu/document/7467408/

[11] J.L. del Val Román, "Industria 4.0: la transformación digital de la industria" Informe CODDii, Conferencia de Directores y Decanos de Ingeniería Informática, Facultad de Ingeniería de la Universidad de Deusto. Bilbao, España. 2016.

Análisis de plataformas de Cloud Computing. Caso Microsoft Azure y Amazon Web Services, haciendo uso de versiones privadas de prueba en entornos educativos

Silvia Edith Arias[1], Laura Mónica Vargas[2,3], Alejandra Di Gionantonio[1], Diego Serrano[1], Adriana Cucchi[1], Paula Sosa[1], Ezequiel Ambrogio[1], Florencia Espeche[1], Pablo Jornet[1], Daniel Arch[1]

s_autn@hotmail.com, {laura.monica.vargas, ing.alejandradg, diegojserrano, adriana.beat, sosa.pau, ezequielambrogio, flor.espeche, pablo.jonet23} @gmail.com, daniel.arch@pjn.gov.ar

[1]Laboratorio de Investigación de Software, Departamento de Ingeniería en Sistemas de Información, Facultad Regional Córdoba, Universidad Tecnológica Nacional

[2] Laboratorio 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 la etapa actual de este proyecto de investigación, abordamos el análisis de dos plataformas de Cloud Computing. Es el caso de Microsoft Azure y Amazon Web Services, haciendo uso de versiones privadas de prueba en entornos educativos.

Teniendo en cuenta sus características principales, se procede a realizar una comparación de las plataformas mencionadas, con la finalidad de elaborar un compendio de parámetros a tener en cuenta a la hora de decidir por una u otra. En aras de cumplir el objetivo de este trabajo que resulta ser el de explorar y obtener conocimientos teóricos sobre marcos de trabajo de plataformas de 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 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, cloud computing.

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 SeCyT – UTN1166, 2010-2011, Resolución 26/10, 2010 SeCyT del Rectorado de UTN y "Marcas de Agua Seguras en Imágenes para identificación del propietario". Proyecto ID promocional. Código SeCyT- UTN EIPRCO753, 2008-2009, Resolución 75/08 SeCyT 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 – Nº 1 (2016).

III. INTRODUCCIÓN

En el marco de Cloud Computing la investigación se centra en el análisis de las características principales de plataformas como Amazon y Microsoft Azure.

En 2006, Amazon Web Services (AWS) comenzó a proporcionar servicios de infraestructura de TI para empresas en forma de servicios web, más conocido hoy como cloud computing. Uno de sus principales beneficios es la oportunidad de reemplazar importantes gastos anticipados en infraestructura con costos variables reducidos que se escalan con su negocio. Gracias a la Cloud Computing, las empresas ya no tienen que planificar ni adquirir servidores y otra infraestructura de TI con semanas o meses de antelación. [1]

Amazon Web Services ofrece herramientas en las siguientes categorías:

- **Cloud computing**: todo lo necesario para la creación de instancias y el mantenimiento o el escalado de las mismas.

- **Bases de datos**: distintos tipos de bases de datos pueden permanecer en la nube mediante el servicio Amazon RDS, que incluye distintos tipos a elegir como MySQL, PosgreSQL, Oracle, SQL Server y Amazon Aurora, o Amazon DynamoDB para NoSQL.

- **Creación de redes virtuales**: permite la creación de redes privadas virtuales a través de la nube, gracias principalmente al servicio Amazon VPC.

- **Aplicaciones empresariales**: Amazon WorkMail es el servicio de correo empresarial que ofrece Amazon, al que pueden unirse otros servicios como Amazon WorkDocs y Amazon WorkSpaces.

- **Almacenamiento y gestores de contenido**: tipos de almacenamiento diferentes, tanto para archivos con acceso regular, poco frecuente o incluso como archivo. Amazon S3 es el servicio principal, aunque complementan la oferta otros como Amazon Glacier o Amazon EBS.

- **Inteligencia de negocios** o Business Intelligence (BI): sistemas para análisis de datos empresariales a gran escala y otros servicios para la gestión de flujos de datos.

- **Gestión de aplicaciones móviles**: herramientas como Amazon Mobile Hub permiten la gestión, creación, testeo y mantenimiento de aplicaciones móviles a través de la nube.

- **Internet de las cosas (Internet of Things):** para establecer conexiones y análisis de todos los dispositivos conectados a Internet y los datos recogidos por los mismos.

- **Herramientas para desarrolladores**: para almacenar código, implementarlo automáticamente o incluso publicar software mediante un sistema de entrega continua.

- **Seguridad y control de acceso**: se pueden establecer autenticaciones en varios pasos para poder proteger el acceso a sus sistemas internos, ya estén en la nube o instalados de forma local en sus instalaciones. [2]

Microsoft Azure es la plataforma de computación en nube pública de Microsoft.

Microsoft Azure se considera ampliamente como un servicio de plataforma como servicio (PaaS) e infraestructura como servicio (IaaS). Microsoft introdujo Azure en octubre de 2008. La plataforma de la nube se llamaba originalmente Windows

Azure, pero fue cambiada de marca a Microsoft Azure en abril de 2014.
Azure es una plataforma de servicios en la nube pública que admite una amplia selección de sistemas operativos, lenguajes de programación, plataformas, herramientas, bases de datos y dispositivos. Puede ejecutar contenedores de Linux con integración de Docker, compilar aplicaciones con JavaScript, Python, .NET, PHP, Java, Node.js y crear back-ends para dispositivos iOS, Android y Windows.[3]
Para garantizar la disponibilidad, Microsoft tiene centros de datos de Azure ubicados en todo el mundo. A partir de enero de 2016, Microsoft dijo que los servicios de Azure están disponibles en 22 regiones de todo el mundo, incluyendo en los Estados Unidos, Europa, Asia, Australia y Brasil.
Microsoft clasifica los servicios de Azure en 11 tipos principales de productos:

- **Cómputo:** Estos servicios proporcionan máquinas virtuales, contenedores, procesamiento por lotes y acceso a aplicaciones remotas.
- **Web y móvil:** Estos servicios soportan el desarrollo y despliegue de aplicaciones web y móviles, y también ofrecen funciones para la administración, notificación y generación de informes de API.

- **Almacenamiento de datos:** Esta categoría incluye las ofertas de base de datos como servicio para SQL y NoSQL, así como almacenamiento en la nube no estructurado y en caché.
- **Analítica:** Estos servicios proporcionan análisis y almacenamiento distribuidos, así como analítica en tiempo real, análisis de Big Data, lagos de datos, aprendizaje automático y Data Warehousing.

- **Redes:** Este grupo incluye redes virtuales, conexiones y pasarelas dedicadas, así como servicios para la administración del tráfico, el equilibrio de carga y el alojamiento del sistema de nombres de dominio (DNS).

- **Red de entrega de contenido y medios (CDN):** Estos servicios incluyen streaming por demanda, codificación y reproducción e indexación de medios.

- **Integración híbrida:** Son servicios para la copia de seguridad de servidores, la recuperación de sitios y la conexión de nubes privadas y públicas.

- **Gestión de identidades y accesos (IAM):** Estas ofertas garantizan que solo los usuarios autorizados puedan utilizar los servicios de Azure, y ayudar a proteger las claves de cifrado y otra información confidencial.

- **Internet de las cosas (IOT):** estos servicios ayudan a los usuarios a capturar, monitorear y analizar los datos de IOT, de sensores y otros dispositivos.

- **Desarrollo:** Estos servicios ayudan a los desarrolladores de aplicaciones a compartir código, probar aplicaciones y rastrear posibles problemas. Azure soporta una variedad de lenguajes de programación de aplicaciones, incluyendo JavaScript, Python, .NET y Node.js.

- **Gestión y seguridad:** Estos productos ayudan a los administradores de nube a gestionar su implementación de Azure, a programar y ejecutar trabajos, y a crear automatización. Este grupo de productos también incluye capacidades para identificar y responder a amenazas de seguridad en la nube.[4]

Diferencias entre Microsoft Azure y Amazon Web Services (AWS)

- AWS está enfocado más en modelos serverless proporcionando herramientas completas para dejar aún lado el hardware, licenciamiento y administración costosa.

- Azure está enfocado en modelos de nubes híbridas con la ventaja de que los modelos on-premise conviven mejor por tener ya

licenciamiento Microsoft. Permite una fácil integración con otras herramientas Microsoft y garantía seguridad con back up multi-nube.[5]

- AWS cuenta con escalabilidad y flexibilidad natural sin mínimos de consumo.

- Azure otorga paquetes de almacenamiento predeterminados.

- AWS cuenta ya con productos para soluciones de machine learning altamente automatizados sugiriendo el comportamiento de los algoritmos.

- Azure cuenta con también con productos para ML solo que menos automatizados, es necesario más desarrollo y con ciertas limitantes para la operación y adaptabilidad.[6]

- AWS proporciona varias capacidades y servicios de seguridad para mejorar la privacidad y controlar el acceso de redes. Entre ellos se incluyen:

1. Los firewalls de red integrados en Amazon VPC y las capacidades de firewall para aplicaciones web existentes en AWS WAF permiten crear redes privadas y controlar el acceso a las instancias y aplicaciones
2. Cifrado en tránsito con TLS en todos los servicios
3. Opciones de conectividad que permiten conexiones privadas o dedicadas desde la oficina o entorno on-premise [7]

- Azure ofrece una amplia gama de opciones de seguridad configurables, así como la capacidad de controlarlas, por lo que puede personalizar la seguridad para satisfacer los requisitos exclusivos de las implementaciones de su organización.[3]

Se pueden añadir opciones de autentificación para el uso de los servicios en la nube o el acceso a la información de la empresa.

Permite la creación de plataformas de nubes privadas para que los datos de sus aplicaciones locales tengan capas de seguridad añadidas. De esta forma le será más fácil tanto gestionar los datos de su empresa como evitar el acceso por parte de personas no autorizadas. [8]
Esta síntesis se muestra en la figura 1.

IV. Objetivos y Líneas de Investigación y Desarrollo

El proyecto se inscribe dentro de los lineamientos de investigación en Seguridad Informática.

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.

En trabajos futuros nos avocaremos al análisis y la investigación de Nubes Públicas, para la posterior comparación con las plataformas privadas ya estudiadas.
Mediante la aplicación del método empírico analítico basado en la experimentación y observación de los servicios prestados por plataformas de Nubes para alojar imágenes médicas con marcas de agua.

V. Materiales y Métodos

Cloud Computing es una plataforma computacional de trabajo que ofrece recursos, tales como infraestructura, aplicaciones, procesamiento para ser consumidos bajo demanda como un servicio más en Internet.

Cloud computing es un modelo de computación que brinda grandes beneficios a sus usuarios

finales como así también a organizaciones públicas y privadas.

Plataformas Características	AWS	AZURE
Servicios de cómputo	Elastic Compute Cloud (EC2)	Virtual Machines (VMs)
App Hosting	Amazon Elastic Beanstalk	Cloud Services Azure Websites and Apps Azure Batch Azure Scheduler Logic Apps
Computación sin servidor	AWS Lambda	Azure Functions
Container Support	EC2 Container Service	Docker Virtual Machine Extension
App/Desktop Services	Amazon WorkSpaces Amazon AppStream	Azure RemoteApp
ALM & Code Editor	AWS CodeDeploy	Azure Visual Studio Online
Escabilidad Automática (auto scaling)	Sí, Amazon CloudWatch.	Sí, Autoscaling application block y Windows Azure Fabric Controler
Blueprints / Imágenes para acelerar el aprovisionamiento	Sí, AMI: Imagen de máquina Amazon.	Sí, Imágenes provistas en una galería y también imágenes propias guardadas.
Soporta Sistema Operativo Windows	Sí, -Windows Server 2003 R2. -Windows Server 2008. -Windows Server 2008 R2. -Windows Server 2012.	Sí, -Windows Server 2012 Data Center. -Windows Server 2008 R2 SP1.
Soporta Sistema Operativo Linux	Sí, -SUSE Linux Enterprise Server. -Red Hat Enterprise Linux.	Sí, -OpenSUSE 12.3. -SUSE Linux Enterprise Server 11 SP2. -Ubuntu Server 12.04 LTS. -Ubuntu Server 12.10. -Ubuntu Server 13.04. -OpenLogic CentOS 6.3. -Ubuntu Server 12.10 DAILY.
Soporte para lenguajes	C++ -C# -Java -Perl -Python -Ruby	-.NET -Java -Node.js -Python
Soporte para almacenamiento de datos	-Amazon S3. -Amazon Relational DB Service. -Amazon SimpleDB. -SQL Server Express. -SQL Web. -SQL Server STD. -Amazon Redshift	-SLQ Relacional. -Almacenes de tablas NoSQL. -Blob no estructurado. -Amazon Dynamo DB
Soporte para Colas	Amazon Simple Queue Service	-Windows Azure Service Bus. -Colas FIFO con protocolos REST, AMQP, WS.
Almacenamiento en caché	Amazon Elastic Cache	Azure Managed Cache (Redis Cache)
Orquestación de datos	AWS Data Pipeline	Azure Data Factory
Almacenamiento híbrido en la nube	AWS Storage Gateway	StorSimple
Copias de seguridad	Amazon Glacier	Azure Backup
Planificación de recuperación de desastres	None	Azure Site Recovery
Red de entrega de contenidos (CDN)	Amazon CloudFront	Azure CDN
Servidor Web	-Apache. -IIS. -Otros	IIS v7.5
Alternativas de Hipervisores	XEN y LXC (Linux Containers)	XEN y LXC (Linux Containers)

Figura 1. Comparación entre Azure y AWS [9] [10]

VI. Formación de Recursos Humanos

Este proyecto está conformado por docentes-investigadores pertenecientes a la carrera de grado de Ingeniería en Sistemas de Información.
El grupo está compuesto por una Directora, cuatro profesores investigadores de apoyo, tres ingenieros aspirantes a incorporarse a la carrera de investigador y dos becarios 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.
Se dirigirán trabajos finales sobre la temática abiertos a estudiantes de Ingeniería en Sistemas de Información.

REFERENCIAS

[1] https://aws.amazon.com/es/about-aws/

[2] https://www.ticportal.es/temas/cloud-computing/amazon-web-services

[3] https://docs.microsoft.com/es-es/azure/security/azure-security

[4] https://searchdatacenter.techtarget.com/es/definicion/Microsoft-Azure-Windows- Azure

[5] https://www.inbest.cloud/comunidad/aw s-vs.-azure-comparativo-general

[6] https://www.bravent.net/migrar-a-la-nube-azure-o-aws

[7] https://aws.amazon.com/es/security/

[8] https://www.ticportal.es/temas/cloud-computing/microsoft-cloud/microsoft-azure/seguridad-azure

[9] https://blog.apterainc.com/amazon-web-services-vs-microsoft-azure-the-real-difference

[10] https://stackify.com/microsoft-azure-vs-amazon-web-services-vs-google-compute-comparison/

Aspectos de Seguridad en Sistemas de Voto Electrónico

Pablo García[1] Silvia Bast[1] Germán Montejano [1 2]

[1]Departamento de Matemática
Facultad 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
[pablogarcia, silviabast]@exactas.unlpam.edu.ar

[2]Departamento de Informática
Facultad 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

RESUMEN

Las discusiones acerca de la aplicación del Voto Electrónico en Argentina se han profundizado en los últimos tiempos. Si bien existen grandes cuestionamientos sobre el uso de la tecnología en los procesos electorales, la idea de desarrollar un sistema de voto electrónico que cumpla con las características de anonimato y transparencia, no parece inviable si se aplican las condiciones de seguridad apropiadas. Es imprescindible que todos los ciudadanos perciban de manera incontrastable la transparencia del proceso; eso implica que cada votante pueda verificar los resultados y a la vez, alcanzar un convencimiento total de que el proceso se llevó a cabo de manera totalmente confiable.

Los investigadores del proyecto "Aspectos de en Seguridad en Proyectos de Software" desarrollan en paralelo dos modelos que pueden aplicarse a los sistemas de voto electrónico:

a. Basado en criptografía homomórfica.
b. Basado en criptografía One Time Pad.

En este trabajo se exponen los avances que se llevaron a cabo para cada uno de los modelos.

Palabras clave: Sistemas de Voto Electrónico, Anonimato, Transparencia, Criptografía Homomórfica, One Time Pad, Verificabilidad E2E.

CONTEXTO

Este trabajo se enmarca en el Proyecto de Investigación: "Aspectos de Seguridad en Proyectos de Software", que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (UNLPam) Resolución N° 488/14 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (FCEyN - UNLPam) y es dirigido por el Doctor Germán Antonio Montejano (Universidad Nacional de San Luis) y codirigido por el Magister Pablo Marcelo García (FCEyN - UNLPam) e incluye a la Magister Silvia Gabriela Bast, al Magister Daniel Vidoret, al Licenciado Adrián García y al Programador Superior Claudio Ponzio como investigadores.

Surge desde la línea de Investigación "Ingeniería de Software y Defensa Cibernética", presentada en [1], 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 Geráis (UFMG, Brasil).

1. INTRODUCCIÓN

a. Modelo Basado en Criptografía Homomórfica

El modelo deberá cumplir con una serie de requisitos [2] que se exigen actualmente a los sistemas de votación electrónica:

- Evidencia física que garantice la transparencia del proceso [3]. En la mayoría de los casos, los nuevos sistemas son híbridos, es decir que producen un voto impreso que se utiliza como garantía de seguridad.
- Utilización de métodos criptográficos cuya seguridad incluya formalidad matemática, tanto en lo referente al anonimato del votante como a la transparencia de los resultados de los comicios. En este modelo se analizan alternativas basadas en esquemas homomórficos como Paillier [4], [5] o ElGamal exponencial [6], [7].
- Aplicación del concepto de independencia del software [8].
- Definir un modelo concreto para la aplicación de verificabilidad "End to End" (E2E) [9], [10].
- Aplicación de técnicas que eviten que un votante pueda demostrar por quién votó.

Los avances realizados en el proyecto durante 2018 se relacionan con tres aspectos:

- Análisis de la criptografía concreta que se va a aplicar en el modelo.
- Estudio de la forma exacta en que deben realizarse las comunicaciones de datos en el modelo de voto electrónico propuesto.
- Definición de una interface apropiada.

b. Modelo basado en One Time Pad (OTP-Vote)

Este modelo teórico se concentra en los aspectos de confidencialidad e integridad de los datos de un sistema de voto electrónico [11] y se basa en la siguiente premisa fundamental: En los sistemas de voto electrónico es necesario proteger:

- De manera indefinida la privacidad del votante, aún después de finalizada la elección, dado que en caso de que algún intruso obtenga una copia digital de registros que permitan relacionar al votante con su voto contará con todo el tiempo para intentar descifrarlo.
- La seguridad de los datos de los votos mientras dure el proceso electoral, dado que luego la información se hace pública.

El modelo propuesto hace uso de:

- **Claves One Time Pad (OTP)** que cumplen con la característica de Secreto Perfecto de Shannon [12] (lo que significa que aún un adversario con potencia de cómputo infinita no puede deducir el texto plano a partir del texto cifrado).
- **El esquema de almacenamiento denominado Múltiples Canales Datos único (MCDU)** y sus fórmulas propuestas para alcanzar dimensiones con comportamiento óptimo [13] y [14], aplicando operación XOR con la redundancia apropiada [15] para el almacenamiento de los datos.

El modelo asegura:

- Anonimato incondicional.

- Seguridad computacional que puede llevarse a cualquier nivel exigible durante el proceso eleccionario.

El proceso eleccionario se lleva a cabo en tres grandes etapas:

- Preparación de la elección.
- Desarrollo de la elección.
- Cierre de la elección y recuento de votos.

Los datos que están incluidos en los procesos son:

- Claves: Claves de las Autoridades (*CA*), Clave del Voto (*CV*) y Clave de Descifrado (*CD*).
- Archivos binarios de datos: Archivo Binario de Votos (*ABV*) y Archivo Binario de Votos Descifrado (*ABVD*)
- Tablas relacionales: IdVotos, Candidatos, Cargos, Atributos, Ubicaciones, Votos Planos

Para que el modelo teórico pueda ser efectivamente implementado, deben resolverse aspectos tales como:

- La seguridad de las *CA* que intervienen en el proceso.
- La inalterabilidad del *CV* que se genera como parte del proceso de desarrollo de la elección, en dos aspectos:
    - su almacenamiento mediante XOR en la *CD*.
    - como aporte a la Contribución Final del Voto que modificará finalmente el *ABV*.
- La confidencialidad e integridad en las tablas relacionales involucradas.
- *ABV* y *CD* deben ser accedidos y modificados sólo por el sistema de la forma propuesta por el Modelo.
- Propuesta del proceso de auditoría.
- Análisis de la semántica de los datos del voto, para incorporar atributos de control y de encriptación, variaciones en los datos almacenados de los votos y profundización de las posibilidades de recuperación semántica de los bits en las tuplas.
- Generación de un mecanismo que, teniendo en cuenta la semántica de los datos de voto (que es configurable y puede variar entre elecciones) facilite la generación de las tablas relacionales resultantes.
- Asegurar la comunicación entre el usuario y el sistema en los momentos en que la misma se produce y la transmisión de datos entre estaciones y servidor.

Una vez que los puntos mencionados estén resueltos, será posible demostrar que el sistema resultante de la investigación es confiable y seguro.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

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 homomórfica, que es la línea a cargo del Magister Pablo Marcelo García.
- Basados en criptografía One Time Pad, que es la línea que lleva adelante la Magister Silvia Gabriela Bast.

3. RESULTADOS Y OBJETIVOS

Los avances del grupo de trabajo que han surgido durante 2018 son:

a. En el ámbito de la criptografía homomórfica, este grupo de trabajo ha realizado las siguientes acciones:

    - Una encuesta online para obtener respuestas de personas de todo nivel (desde expertos informáticos hasta votantes comunes) para recabar opiniones sobre la forma exacta que debería tener la

interface de un sistema de voto electrónico.

- Se incorporan al proyecto dos especialistas específicos en comunicaciones de datos para proporcionar metodologías de transmisión de datos que garanticen los niveles de seguridad exigibles. Los mismos se encuentran en proceso de investigación para proporcionar un modelo de comunicación que cumpla con los requisitos.

A futuro, se pretende llevar a cabo las siguientes acciones:

- Aplicar la interface seleccionada al modelo a implementar.
- Desarrollar el modelo de transmisión de datos e implementarlo en la aplicación final.
- Realizar un permanente relevamiento de aplicaciones orientadas al voto electrónico, que permita detectar falencias y proponer mejoras en el nuevo modelo.

b. Con respecto a la línea de OTP Vote se ha avanzado sobre:

- El estudio de la semántica de los datos del voto, generando una propuesta de almacenamiento de los datos, que incluye atributos de control y de encriptación, variaciones en los datos almacenados de los votos, que apunta a oscurecer el contenido de los mismos para otorgar así, mayor seguridad al proceso eleccionario.
- El refinamiento de los procesos de recuperación y generación de votos planos [16]. Éste último mismo incluye un parser que, dada la configuración de los datos de la elección, que se establecen en la etapa inicial del proceso, genera la tabla de Votos Planos a medida que va realizando la recuperación de los sufragios emitidos por los electores.

Se continuará trabajando dentro de esta línea de investigación con aspectos tales como:

- Aseguramiento de la comunicación entre usuario-sistema y sistema-servidor de datos
- Propuesta de modelo de auditoría de terceros.
- Automatización del proceso de generación de las tablas relacionales.
- Propuesta de Verificabilidad End to End.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco del presente proyecto se presentan los siguientes puntos relacionados con la formación de recursos humanos:

- Pablo García y Silvia Bast completaron 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 asistió a un encuentro de tutoría con sus directores de Tesis Doctoral el Dr. Germán Montejano (UNSL) y Jeroen van de Graaf, PhD. (UFMG), en el ámbito de la FCFMyN (UNSL).
- Pablo García y Silvia Bast presentaron las modificaciones oportunamente exigidas a 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). Las mismas se encuentran en proceso de evaluación.

5. BIBLIOGRAFÍA

[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 WICC. Ps 769-773. ISBN: 9789872817961. 2013.http://sedici.unlp.edu.ar/handle/10915/27537

[2] **Hao, F, Ryan P.:** "Real -World Electronic Voting. Design, Analysis And Deployment". Cr Press. ISBN-13: 978- 1498714693. ISBN-10: 1498714692. 2017.

[3] **Prince, A.:** "Consideraciones, Aportes y Experiencias para el Voto Electrónico en Argentina". Editorial Dunken. ISBN: 978-987-02-1732-9. 2006.

[4] **Volkhausen T.**: "Paillier Cryptosystem: A Mathematical Introduction". 2006.

[5] **O'Keeffe M.**: "The Paillier Cryptosystem: A Look Into The Cryptosystem And Its Potential Application". The College of New Jersey Mathematics Department. 2008.

[6] **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.

[7] **Koscielny C.**: "A New Approach to the Elgamal Encryption Scheme". Academy of Management of Legnica, Faculty of Computer Science. 2004.

[8] **Rivest R**.: "On the notion of 'software independence' in voting systems". Philosophical Transactions of The Royal Society A, 366(1881):3759–3767. 2008.

[9] **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 EVote-ID 2017.

[10] **Kelsey J., Regenscheid A., Moran T., Chaum D.**: "Attacking Paper-Based E2E Voting Systems". In: Chaum D. et al. (eds).

[11] **Bast S.:** "Optimización de la Integridad de Datos en Sistemas de E-Voting". Tesis de Maestría defendida en la Universidad Nacional de San Luis. 14 de diciembre de 2016. San Luis, Argentina.

[12] **Shannon, C.:** "Communication Theory of Secrecy Systems" - Bell System Technical Journal - 1949.

[13] **García, P.**: "Una Optimización para el Protocolo Non Interactive Dining Cryptographers" - Editorial Académica Española (https://www.eae-publishing.com/ - ISBN-13: 978-3-639-85270-7. ISBN-10: 3639852702. EAN: 9783639852707 – 2017.

[14] **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 (JAIIO, ISSN: 1850-2776). Workshop de Seguridad Informática (WSegI 2013, ISSN: 2313-9110). Páginas 29 a 43. Disponible en: http://42jaiio.sadio.org.ar/proceedings/simposios/Trabajos/WSegI/03.pdf. 2013.

[15] **García P., Montejano G., Bast S., Fritz E.:** "Codificación de Sufragios con Detección de Colisiones en NIDC con Canales Paralelos de Slots" Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2016.

[16] **Bast S., García P., Montejano G.:** "Refinamiento de los Procesos de Recuperación de Sufragios y Generación de Votos Planos en OTP-Vote" Congreso Nacional de Ingeniería en Informática / Sistemas de Información. CoNaIISI 2018.Universidad CAECE – Mar del Plata, Buenos Aires, Argentina. Publicación on line - ISSN 2347-0372

Ciberseguridad en los Sistemas de Control Industrial: Clave para la Ciberdefensa de las Infraestructuras Críticas

Jorge Kamlofsky[1], Samira Abdel Masih[1], Hugo Colombo[1], Claudio Milio[1] y Pedro Hecht[2]

[1] CAETI - Universidad Abierta Interamericana
Av. Montes de Oca 725 – Buenos Aires – Argentina
{Jorge.Kamlofsky, Samira.Abdel Masih, Hugo.Colombo, Claudio.Milio}@uai.edu.ar

[2] Universidad de Buenos Aires, Facultades de Ciencias Económicas, Ciencias Exactas y Naturales e Ingeniería. Maestría en Seguridad Informática, Buenos Aires, Argentina
phecht@dc.uba.ar

Resumen

La automatización de los procesos industriales se realiza mediante los sistemas de control industrial. Con lógica determinista, se diseñaron para funcionar con alta disponibilidad. Por ser robustos y efectivos, se los utiliza para la automatización de los procesos de las infraestructuras críticas: plantas de generación y distribución de energía, potabilizadoras de agua, sistemas de semaforización, entre otros.

Tradicionalmente su seguridad se basó en el aislamiento físico y en las enormes diferencias tecnológicas con los sistemas informáticos. Su interconexión con las redes administrativas y con Internet brindó mayor flexibilidad y eficiencia, pero quedaron expuestos a una gran cantidad de vulnerabilidades y amenazas.

En este proyecto se estudian las vulnerabilidades de estos sistemas y se proponen soluciones basadas en mejoras de procesos, comunicaciones y en criptografía e inteligencia artificial.

Palabras clave: seguridad en redes industriales, seguridad en scada, criptografía compacta, ciberdefensa en infraestructuras críticas.

Contexto

Los proyectos radicados en el CAETI[1] se clasifican en tres líneas de investigación. Este proyecto se enmarca dentro la línea de Automatización y Robótica. Se inició en Abril de 2014.

Introducción

La producción industrial a gran escala se automatiza mediante los Sistemas de Control Industrial (ICS, de sus siglas en inglés). Son sistemas de tele-mando y tele-control de procesos compuestos por autómatas industriales (según sus siglas en inglés): RTU (Remote Terminal Unit), PLC (Programmable Logic Controller), DCS (Distributed Control System) y/o PAC (Programmable Automation Controller) que pueden interconectarse [1]. Se les conectan entradas y salidas, discretas y/o analógicas como ser: micro-switches, sensores de temperatura, actuadores para encendido de motores, llaves, etc. Poseen procesadores de pequeño porte. Su lógica es determinista,

[1] CAETI: Centro de Altos Estudios en Tecnología informática, UAI.

lo cual favorece a la alta disponibilidad, esencial en el ambiente industrial [2].

Los ICS son monitoreados desde una Interfaz Hombre-Máquina (en inglés, HMI: Human Machine Interface). Se supervisan y controlan en tiempo real desde sistemas informáticos llamados SCADA (del inglés: Supervisory Control and Data Acquisition). Los SCADA suelen incluir terminales HMI visuales con pantallas táctiles, terminales para mantenimiento e ingeniería [2]. Los ICS son muy robustos, y por ello automatizan procesos que requieren uso continuo: plantas de potabilización de agua, producción y distribución de energía, transporte, siderúrgicas, entre otras. Es decir, están en las infraestructuras críticas de naciones. El aislamiento físico y las diferencias con la tecnología informática les dieron una falsa sensación de seguridad por ocultamiento [3, 4].

En búsqueda de mayor flexibilidad y eficiencia, los desarrollos en tecnologías de la información y telecomunicaciones, incentivaron a interconectar los SCADA con las redes corporativas e incluso, a Internet. Así, los ICS quedaron expuestos a amenazas y riesgos que suponen serias consecuencias [5]. En 2010, el sistema SCADA de una planta de enriquecimiento de uranio de Irán fue atacada por un virus llamado Stuxnet [6]. La comunidad internacional mostró gran preocupación por la seguridad de las infraestructuras basadas en estas tecnologías [7-9] y trabaja en soluciones [10-12].

En el ámbito de las tecnologías informáticas se tiene amplia experiencia en Seguridad. Las sugerencias de las normas ISO 27000 y NIST SP800 [13,14] y una gran cantidad de soluciones técnicas, ayudan a proteger a los activos informáticos. En el ámbito industrial, las recomendaciones de la norma NIST SP800-82 [15] y el análisis de vulnerabilidades presentado en [20] son de gran ayuda. Sin embargo, la disponibilidad de soluciones es escasa. Para asegurar un sistema informático, se mencionan en [13] tres pilares básicos: disponibilidad, confidencialidad e integridad. Las soluciones para los ICS, deberían enfocarse, entonces, en mejorar la confidencialidad y/o integridad.

En este proyecto se analizan diferentes recomendaciones de seguridad y se proponen soluciones basadas en mejoras en los procesos y en las tecnologías y topologías de las redes. Se desarrollan soluciones que mejoren confidencialidad e integridad basadas en criptografía e Inteligencia artificial, cuya aplicación en estos sistemas novedosa.

Líneas de Investigación, Desarrollo e Innovación

En el proyecto se identifican tres líneas de investigación:

La primer línea de investigación se denomina Gestión de la Seguridad Informática. El punto de partida fue el análisis de normas [13-15]. Se extiende al análisis de técnicas y procesos de seguridad, topologías de redes, redes industriales, análisis de malware, hacking e informática forense, entre otros.

La segunda línea se denomina Inteligencia Artificial (IA) Aplicada a los ICS. Se dedica al estudio de técnicas de IA con la intención de implementarlas de diversas formas en los ICS.

La tercer línea se denomina Criptografía Aplicada a los ICS, trata la aplicación de algoritmia criptográfica clásica y criptografía post-cuántica basada en álgebra no conmutativa. Implementar criptografía en el interior del ICS es un novedoso desafío que permitiría de asegurar Confidencialidad en las comunicaciones. Y esto pareciera viable teniendo en cuenta que el criptosistema HK17 [16] es simple y compacto lo cual

permitiría su uso en estos sistemas sin comprometer la disponibilidad.

Resultados y Objetivos

En el marco del proyecto se han obtenido los siguientes resultados:

- **Convenios:** Convenio General de Colaboración entre la Universidad Abierta Interamericana (UAI) y la Universidad Fasta (UFASTA). Se firmó el 5/12/17. En 2018 se realizaron charlas de extensión en cada institución, articuladas desde UFASTA por el InFo-Lab y desde la UAI por el Caeti. Su objetivo es reforzar los conocimientos en Ciberseguridad de sus investigadores.
- **Dictado de Cursos, Charlas y Seminarios:** Taller de Criptografía (UAI, 27/9/17 y UFASTA 12/7/18 y 29/10/18). Conversatorio acerca de Alain Turing (UNGS, 10/11/17), Charlas Ciberseguridad para Gente Común (UAI, 23/9/16 y E.C. N°22, 4/9/18).
- **Extensión:** Jornadas de Informática Forense dictadas por los investigadores de UFASTA Juan Alberdi y Bruno Constanzo (UAI, 16/4/18 y 27/8/18).
- **Participación en Talleres de Ciberseguridad del MinSeg:** Hacia la construcción de un campo de I+D+i en Seguridad (MinCyT, 15/12/16), Taller Británico / Argentino: Tendencias Globales e Innovación en Ciberseguridad. Experiencias en Argentina y Gran Bretaña (MinSeg, 19/3/18 al 21/3/18).
- **Tesis:** Pablo Oviedo y defendió su tesis de Licenciatura en Matemática denominada: Fundamentos matemáticos de computación cuántica en el algoritmo de Shor, para la factorización prima de números enteros (UAI, 2017).
- **Patentes:** Res. de Derechos de Autor del algoritmo criptográfico HK17. RE-2017-23378132-APN-DNDA (6/10/17).
- **Concurso:** Presentación y aceptación del algoritmo HK17 [16] en CFP de la NIST [17] (USA, 30/11/17).
- **Publicaciones:** Publicación de siete trabajos con referato [18, 19, 21-25]. De ellos: [21, 23, 24] fueron premiados en los respectivos congresos. Los trabajos [8, 26-30] han citado a alguno de ellos.
- **Desarrollos:** Se desarrolló un ICS portable para demo y experimentación y se implementó un Honeypot industrial.

El objetivo general del proyecto es crear soluciones de seguridad para los ICS.

Los objetivos particulares más destacados son: aplicar modelos de IA y criptografía en los ICS y mejorar procesos y topologías de red para asegurar la conexión del ICS/SCADA a la red corporativa y a Internet. Algunos objetivos específicos son: solucionar los ataques presentados al algoritmo HK17 [19, 31], desarrollar modelos de IA y criptografía para detectar conexiones anómalas y ataques a la red.

Objetivos subyacentes: difundir conocimientos que ayuden al abordaje del problema de estudio y además, mejorar la participación de los alumnos en los resultados del proyecto.

Formación de Recursos Humanos

El proyecto está dirigido por el Esp. Lic. Jorge Kamlofsky quien está cursando un Doctorado. Los resultados colaborarán con el desarrollo de su Tesis Doctoral. Está co-dirigido por la Dra. Samira Abdel Masih quien se enfoca en las tutorías de Tesis de Licenciatura en Matemática. En el proyecto colabora el Dr. Pedro Hecht. Integran el proyecto el PhD. Hugo Colombo y el Ing. Claudio Milio, docentes de la Facultad de Tecnología Informática de la UAI quienes adquieren nuevos conocimientos de ciberseguridad.

El equipo de investigación se completa con alumnos de la Facultad de Tecnología Informática de la UAI: Nicolás Carella, Federico Tabarez Rosa, Daniel Sola, Alejandro Vassallo Lagos, Sebastian Caldarola, Ricardo Goffi, Juan Di Modugno, Jonatan Schmidt, Valeria Galván, Federico Arango, Federico Romero, Lucas Barros, Nahuel Perez, Enrique Belaustegui, Angel Orlauskas, Nicolás Mayer, Fernanda Lopez y Cecilia Prieto. Se desempeñan como auxiliares de investigación. El conocimiento adquirido se incorporará en sus trabajos finales de carrera. Enrique Belaustegui es un alumno avanzado de Ingeniería en Sistemas con experiencias con ICS, tiene un rol destacado en el proyecto. Fernanda Lopez y Cecilia Prieto están realizando sus Tesis de final de Licenciatura en Matemática tratando temas de criptografía poscuántica. Los alumnos Juan Perdiguizzi y Yahel Bayarsky se encuentran realizando sus prácticas en el marco de la materia Práctica Profesional Supervisada de la carrera Ingeniería en Sistemas. Juan Perdiguizzi logró el desarrollo de un ICS portable, para experimentación y demostración. Yahel Barbasky está elaborando un documento que resume las normas de seguridad de la información [13, 14].

En el proyecto también participan alumnos de otras instituciones: Pablo Pizio y Nicolás Ferella están iniciando su tesina de Licenciatura en Ciencias de la Computación de la Universidad Nacional de La Plata acerca de Criptografía poscuántica. Juan Ferreyra es alumno del Instituto IESI Itemed, y estudia aspectos de la Informática Forense.

El problema que se estudia en este proyecto se encuentra latente en las infraestructuras críticas e industriales del mundo. Resulta muy atractivo tanto para docentes como para estudiantes.

Referencias

[1] Miguel. “¿DCS, PLC, PAC o RTU?,” Control Real Español, (2015). Disponible en: https: // controlreal.com / es / dcs - o - plc – o – pac – o - rtu/. [Consultado: 8/03/2017].

[2] Romero Mestre, H. “Çiberseguridad en sistemas de control industrial o ICSs.” Trabajo Final de Master. Incibe, UOC, URB, Universitat Autonoma de Barcelona, (2018).

[3] Courtois, N. "The dark side of security by obscurity, and Cloning MiFare Classic Rail and Building Passes Anywhere, Anytime." IACR Cryptology ePrint. 137, (2009).

[4] Menezes, A., Van Oorschot, P., Vanstone, S. "Handbook of applied cryptography”. CRC press, (1996).

[5] Sanchez, P. "Sistema de Gestión de la Ciberseguridad Industrial." Trabajo Final de Master. Univ. Oviedo, (2013).

[6] Englert, M. ”Cyber meets nuclear Stuxnet and the cyberattacks on Iranian centrifuages." Deutschen Physikalischen Gesellschaft, (2013).

[7] Corvalan, F. "Seguridad de Infraestructuras Críticas: Visión desde la Ciberdefensa." III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, (2015).

[8] CEEAG "La Ciberguerra. Sus Impactos y Desafıos." Centro de Estudios Estratégicos de la Academia de Guerra, Ejército de Chile, (2018).

[9] Andreeva, O., Gordeychik, S., Gritsai, G., Kochetova, O., Potseluevskaya, E., Sidorov, S. and Timorin, A. Industrial control systems vulnerabilities statisttics." Kaspersly Labs, (2016).

[10] Sajid, N., Patel, S. and Patel, D. ”Assessing and augmenting SCADA cybersecurity: A survey of techniques." Computers and Security 70, (2017): 436-454.

[11] Blackmer, M. ”Cibersecurity for Industrial Control Networks." III Conferencia Internacional y Taller de Ciberseguridad e Infraestructuras Críticas de Información, Buenos Aires, 2015.

[12] Simoes, P., Cruz, T., Proenca, J. and Monteiro, E. "Honeypots especializados para Redes de Control Industrial." VII Congreso Iberoamericano de Seguridad Informática, Panamá, 2013.
[13] ISOTools. ”ISO 27001." (2015). Disponible en: https:// www.isotools.org / normas / riesgos-y-seguridad / iso-27001. [Consultado: 20/01/2019].
[14] NIST. "Special Publication 800 - 30, revision 1." Information Security. National Institute of Standards and Technology, U.S. Departament of Commerce, (2012).
[15] Stouffer, K., Lightman, S., Pillitteri, V., Abrams, M. and Hahn, A. "Guide to Industrial Control Systems (ICS) Security." NIST. Special Publication 800 / 82, revision 2. U.S. Departament of Commerce, (2015).
[16] Hetch, P. and Kamlofsky, J. "HK17: Post Quantum Key Exchange Protocol Based on Hypercomplex Numbers." NIST: National Institute of Standards and Technology, U.S. Departament of Commerce, Post Quantum Cryptography Project, (2017). Disponible en: https: // csrc.nist.gov / CSRC / media / Projects / Post-Quantum-Cryptography / documents / round-1 / submissions / HK17.zip [Consultado: 20/01/2019].
[17] Chen, L., Moody, D. And Liu, Y. "Post Quantum Cryptography, Call for Proposal." NIST: National Institute of Standards and Technology, U.S. Departament of Commerce, Post Quantum Cryptography Project, (2017). Disponible en: https: // csrc.nist.gov / Projects / Post-Quantum-Cryptography / Post-Quantum-Cryptography-Standardization / Call-for-Proposals. [Consultado: 20/01/2019].
[18] Kamlofsky, J., Hetch, P., Izzi, O. and Abdel Masih, S. "A Diffie Hellman compact model over conmutative rings using quaternions." VIII Congreso Iberoamericano de Seguridad Informática, Quito, 2015.
[19] Kamlofsky, J., Colombo, H., Sliafertas, M. y Pedernera, J. "Un Enfoque para Disminuir los Efectos de los Ciber-ataques a las Infraestructuras Críticas." III Congreso Nacional de Ingeniería Informática / Sistemas de Información (CONAIISI 2015), ISSN: 2346-9927. (2015).
[20] INL "Vulnerability Analysis of Energy Delivery Control Systems." Idaho National Laboratory, INL/EXT-10-18381, (2011).
[21] Kamlofsky, J. "Improving a Compact Cipher Based on Non Commutative Rings of Quaternions." XXII Congreso Argentino de Ciencias de la Computación, San Luis, 2016.
[22] Kamlofsky, J., Hecht, P. and Abdel Masih, S. "Post-Quantum Cryptography: An Elementary and Compact Key Exchange Scheme Based on Octonions." IX Congreso Iberoamericano de Seguridad Informática, Buenos Aires, 2017.
[23] Kamlofsky, J. and Hecht, P. "Post-Quantum Cryptography Using Hyper-Complex Numbers." XXIII Congreso Argentino de Ciencias de la Computación, La Plata, 2017.
[24] Kamlofsky, J. and Mieres, J. "A Graph Approach to Improve Crimeware Analysis and Classification." IX Congreso Iberoamericano de Seguridad Informática, Buenos Aires, 2017.
[25] Kamlofsky, J., Abdel Masih, S., Colombo, H., Veiga, D., Costa, E., Milio, C., Semeria, M. y Hecht, P. "Seguridad en las Redes Industriales: Clave para la Ciberdefensa de las Infraestructuras Críticas." XIX Workshop de Investigadores en Ciencias de la Computación, Buenos Aires, 2017.
[26] Castro Lechtaler, A., Cipriano, M., Garcia, E., Liporace, J., Maiorano, A., Malvacio, E., Tapia, N., Dulio, N. y Perez, P. "Secuencias Seudoaleatorias para Criptología." XVIII Workshop de Investigadores en Ciencias de la Computación, Concordia, 2017.
[27] Hecht, J. "Post-Quantum Cryptography: S381 Cyclic Subgroup of High Order." ArXiv preprint arXiv:1704.07238, (2017).
[28] Durcheva, M. and Karailiev, K. "New public key cryptosystem based on quaternions." AIP Conference Proceedings 1910, 060014, (2017).
[29] Arzzolini, C. "Ciberseguridad en la República Argentina y su Perspectiva Futura." Trabajo Final Integrador, Instituto de Inteligencia de las Fuerzas Armadas, 2018.
[30] Remache Rubio, E. "Modelo para la Mitigación de Vulnerabilidades Informáticas en los Servicios Web." Trabajo Final de

Maestría. Pontificia Universidad Católica del Ecuador, 2018.
[31] Li, H., et al. "Cryptanalysis of HK17." IACR Cryptology ePrint Archive 2017: 1259}, (2017).

Clasificando ransomwares para el desarrollo de un detector de código malicioso en ejecución

Iris N. Gastañaga, Fabián A. Gibellini, Pablo S. Frías, Analía L. Ruhl, Leonardo R. Cíceri, Germán N. Parisi, Federico J. Bertola, Paula B. Olmedo, Milagros N. Zea Cárdenas.

Laboratorio de Sistemas / Dpto. de Ingeniería en Sistemas de Información/
Universidad Tecnológica Nacional / Facultad Regional Córdoba
Cruz Roja S/N, 5016

irisg@ciec.com.ar, fgibellini@bbs.frc.utn.edu.ar, pablosfrias@gmail.com, analialorenaruhl@gmail.com, leocic@bbs.frc.utn.edu.ar, germannparisi@gmail.com, federicojbertola@gmail.com, paulabeatrizolmedo@gmail.com, milyzc@gmail.com.

Resumen

En los tiempos actuales, un tipo de software malicioso denominado ransomware, está afectando a muchos usuarios en todo el mundo. Si bien se ha escuchado más sobre aquellos ransomware que cifran archivos del sistema de la computadora infectada y luego piden una recompensa por la clave necesaria para descifrar los archivos, existen múltiples variedades de ransomwares. La recompensa que se pide, generalmente es monetaria y virtual, dentro de las cuales la más conocida es el Bitcoin.

En el presente documento se refuerza y amplía la categorización de ransomwares presentada por Sanggeun et al., buscando comprenderlos para seleccionar un conjunto de estas categorías que permitan definir parámetros e indicadores para desarrollar un detector de ransomwares en ejecución.

Palabras clave: Ransomwares, Cryptoransomwares, Cifrado, Algoritmos.

Contexto

Los resultados presentados en este documento forman parte del Proyecto Homologado "Sistema de Detección de Código malicioso - ransomware", cuyo código es SIUTNCO0004991. El mismo se lleva a cabo en la UTN - FRC y cuyo objetivo general es "Desarrollar un sistema de detección del malware ransomware durante su ejecución en una computadora víctima y detener su avance antes de cifrar todos los archivos, basado en metodologías abiertas para identificar y analizar sus vulnerabilidades".

Introducción

Desde finales de la década de los 80 encontramos métodos de ciber-extorsión, especialmente con la llegada del gusano Morris en el año 1988 [1], el primer malware autorreplicable, que afectó a miles de servidores. En el 2004 Young y Yung indicaron, que los futuros ataques, resultarían de una fuerte combinación entre criptografía y malware para atacar los sistemas. En ese año estaba emergiendo una nueva forma de malware en el ciberespacio, conocida como ransomware, que comenzaba a llamar la atención entre investigadores y practicantes de la seguridad informática de sistemas e imponía graves amenazas a la protección de los activos de información [2].

Un ransomware es una forma de software malicioso utilizado en ataques en los que, no se busca destruir irreversiblemente información, sino cifrarla y cobrar por el servicio de recuperación de los datos cifrados [3] [4].

La forma más común de pagar rescates de ataques con ransomwares es a través de criptomonedas, la más conocida y utilizada es el Bitcoin (BTC) [2]. La plataforma sobre la que se maneja el Bitcoin hace que sean

imposible rastrear quién la ha realizado [5]. Este modelo parece ser rentable para las organizaciones criminales que orquestan los ataques sacando provecho a las criptomonedas, asegurándose privacidad al momento de recibir la transferencia de los rescates.

Dentro de los ataques cibernéticos ocurridos antes del 2017 [6] [7] se mencionan el CryptXXX [8], Petya [9], Locky ransomware y Cerber.

En el caso de Cerber, este es conocido como un RaaS (Ransomware as a Service). Un RaaS es un ransomware, que posee parámetros configurables y es ofrecido en la llamada Deep Web a cambio de dinero. Entre los parámetros a configurar de Cerber se encuentra si el origen de la computadora víctima es atacable o no según una lista de países que no deben ser atacadas [10].

Otro ejemplo de RaaS es Philadelphia, al que se le podía configurar la nota que se mostraba en la pantalla de la víctima, detallando la cantidad de bitcoins pedidos por el "rescate" de los archivos y la cantidad de veces que tenía que confirmar la víctima el pago del rescate [11].

Durante el 2017 aparecieron nuevos y múltiples variaciones de los ransomwares ya existentes, entre los que se puede mencionar: Cryptoblock, Sage, WannaCry y NotPetya. Un ejemplo notable fue WannaCry, lanzado en mayo y que se expandió rápidamente a nivel mundial. Hasta el momento es el mayor ataque de ransomware mundialmente hablando y es por esto que consideramos el año 2017 como un punto de inflexion para los ataques de ransomwares. Aquellos equipos que no poseían el parche de actualización, brindado en marzo por Microsoft eran vulnerables a WannaCry, inclusive se crearon parches para versiones que se encontraban fuera de soporte, como Windows XP debido a la magnitude del ataque.

Para empezar a comprender este ransomware hay que empezar por entender las "piezas" que lo componen. WannaCry explota la vulnerabilidad CVE-2017-0144 (CVE, 2017) [10], haciendo uso del exploit EternalBlue, junto con la herramienta DoublePulsar. EternalBlue es un script creado por la National Security Agency (NSA) y liberado por un grupo conocido como Shadow Brokers que permite explotar la vulnerabilidad CVE-2017-0144 de los bloques de mensajes del servidor (SMBv1) [14] y ejecutar código remotamente [11]. DoublePulsar es una puerta trasera para Windows a través de la cual se puede subir archivos al sistema donde está instalado. Cuando Wannacry inicia su propagación a una máquina víctima no infectada, primero verifica que ésta tenga instalado DoublePulsar, si no se encuentra instalado utiliza el exploit EternalBlue para descargarlo e instalarlo. Una vez instalado DoublePulsar inicia la descarga del ransomware en la máquina víctima [12].

NotPetya es una variante del ya conocido ransomware Petya que copia el modus operandi de WannaCry explotando la vulnerabilidad EternalBlue. A pesar del potencial daño que podría haber causado, varios investigadores llegaron a la conclusión de que realmente no cifraba los archivos, como inicialmente se creía.

Según Sanggeun et al, se puede clasificar los ransomware según su funcionalidad [13]:

Scareware: alerta falsamente al usuario sobre la presencia de malware en el equipo, ofreciéndole la posibilidad de comprar un falso antivirus para la eliminación de dicho malware. Puede ser distribuido mediante correos electrónicos, ventanas emergentes o alertas internas cuando se navega por internet.

Locker-ransomware: bloquea el acceso al sistema operativo del equipo, de modo que el usuario no tiene posibilidad de usar la computadora, ni siquiera con sucesivos reinicios. Muestran una advertencia indicando que el usuario ha cometido un delito y debe abonar una multa para recuperar el normal funcionamiento del equipo.

Cryptoransomware: escanea el sistema de archivos y/o las unidades de red del equipo con el fin de cifrar el contenido (suele centrarse en las extensiones más comunes, como por ejemplo jpg, txt o mp3). Naturalmente, el acceso al mismo queda inhabilitado. Actualmente los ransomware de este tipo, exigen el pago en bitcoins.

corresponden a la clasificación dada por Sanggeun [13]:

- Scarewares.
- Locker-ransomwares.
- Cryptoransomwares: para esto utilizan algoritmos de cifrado simétrico, asimétrico o ambos. Existen casos en los que utilizan algoritmos de curva elíptica.

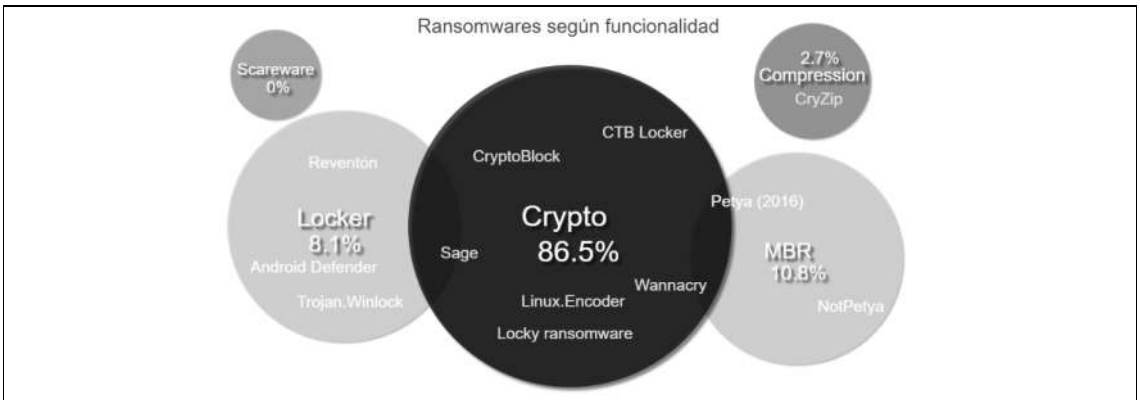


Fig. 1.
Ransomwares según funcionalidad con los representativos en cada categoría.

Partiendo de esta clasificación y estudiando los distintos ransomwares que aparecieron hasta el 2017, se definieron los siguientes objetivos:

- Verificar que la clasificación dada por Sanggeun contemple todos los ransomwares conocidos hasta el momento del estudio.
- Identificar los algoritmos de cifrados más utilizados por los cryptoransomware.
- Identificar una/s categoría predominante para enfocarse en la necesidad de detener estos ataques.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se inscribe dentro de la línea de Seguridad Informática, más precisamente dentro defensas contra código malicioso ransomware.

Resultados y Objetivos

Para la clasificación de ransomwares se tuvo en cuenta el modo de ataque y cómo funcionan, agrupándolos en familias o categorías. Las categorías son listadas a continuación, donde las tres primeras se

Borra los archivos originales.

- MBR ransomwares: consiguen permisos de administrador de modo que sea posible modificar el contenido de la Master Boot Record y así impedir el acceso al sistema. Pueden ser clasificado también como un caso específico de Locker-ransomware.
- Compression ransomwares: impide el acceso a los archivos de usuario, mediante la compresión de estos en ficheros cifrados y protegidos con contraseña. Borra los archivos originales.

Para arrivar a conclusions feacientes se realizó un análisis teórico a una muestra aleatoria de 41 ransomwares identificando a qué categoría pertenecen, su característica más distintiva, el año en que se los detectó, sistema operativo sobre el que se ejecuta y su modo de propagación. Se descartaron aquellos que no fueron considerados como ransomwares modernos y herramientas que sirven para la creación de ransomwares, entre estos One Half y Citadel. Quedando un total de 37 ransomwares.

Los resultados de la Fig. 1 expusieron cierta predominancia de los cryptoransomwares sobre el resto, además se puede observar que el porcentaje total

sobrepasa el 100%, esto se debe a que hay ransomwares que se los consideró incluidos en más de una categoría, como por ejemplo Petya del 2016 es tanto de tipo cryptoransomware y MBR.

El siguiente paso, fue clasificar los algoritmos de cifrado que utilizan los cryptoransomwares. De un total de 37 ransomwares, 32 resultaron cryptoransomwares.

En los resultados de la Tabla 1, se observa que los cryptoransomwares usan tanto algoritmos de cifrado simétrico como asimétrico y de hecho hay algunos que usan ambos tipos de cifrado. Es por esto que, sumando los totales, este valor da 39 ransomwares, superando el total de ransomwares analizados.

Dentro de los que utilizan un algoritmo asimétrico, el más usado es el RSA 2048, mientras que, el AES CBC es el más usado dentro de los algoritmos simétricos.

dependiendo de los nuevos ransomwares que vayan surgiendo posteriormente.

En cuanto al segundo objetivo, los algoritmos asimétricos son los más utilizados porque son más difíciles de "romper", lo que genera mayor dificultad en los intentos para desencriptar los archivos cifrados por un ransomware.

Si bien actualmente los algoritmos más usados son los RSA, se ha comenzado a utilizar algoritmos de curva elíptica que son intrínsicamente más complejos. Otro punto importante respecto a los algoritmos, es que para lograr ransomwares más eficientes, es decir con una encriptación más rápida y al mismo tiempo difícil de romper, se están usando los dos tipos de algoritmo, el simétrico para cifrar los archivos y el asimétrico para cifrar la llave del algoritmo simétrico.

Los resultados ratifican la predominancia de los cryptoransomwares, lo que manifestó la

Algoritmos de de cifrado					
Cifrado asimétrico		Cifrado simétrico			
Algroritmos	Cantidad	Algoritmos	Cantidad	Tamaño de bloque	Cantidad
RSA 56	1	#XOR	1		
RSA 512	1	AES CBC	13	256	5
RSA 2048	5			128	5
RSA 4096	1	AES ECB	1	256	1
Curva elíptica	2			128	0
No idenficados	8	Salsa20	2		
TOTAL	18	No idenficados	4		
		TOTAL	21		

Tabla 1.
Clasificación de algoritmos de cifrado utilizados por cryptoransomwares.

Se inició este proyecto en el contexto del ataque cibernético mundial producido por el Wannacry, por lo cual aludir a la palabra "ransomware" era una asociación directa al cifrado masivo de datos.

Respecto a los objetivos planteado como partida, se revalidó la clasificación de Sanggeun [13] y se agregaron dos categorías, los ransomware de tipo MBR y de Compression. Esta clasificación no se la considera absoluta y está sujeta a cambios, tanto en categorías como subcategorías,

necesidad de enfocarse en descubrir una forma de detectar cryptoransomwares mientras se ejecuten en una computadora víctima de manera que sirva como última barrera de defensa durante un ataque y se minimicen la cantidad de arhivos cifrados.

Si bien fue un estudio teórico, este ha cumplido un propósito mayor, dando una clara descripción de los ransomwares ya conocidos en el mundo cibernético y su evolución, clasificándolos y exponiendo la familia de ransomwares en la que los atacantes se han enfocado.

Basado en estos datos se puede comenzar a trabajar de manera detallada sobre aquellos Malwares de tipo ransomware que sean compatibles en este caso con algoritmos tales como AES (algoritmos simétricos) y RSA (algoritmos asimétricos) preferentemente RSA 2048.

Formación de Recursos Humanos

El grupo está compuesto, ademas de Director, Co-Director, investigadores de apoyo, profesores aspirantes a incorporarse a la carrera de investigador, técnicos de soporte, estudiantes investigadores de la carrera de Ingeniería en Sistemas de Información y becarios que forman parte del equipo. Este proyecto contribuirá a la formación y crecimiento de la carrera de los integrantes del mismo. En el caso de los estudiantes y algunos integrantes se iniciarán en la línea de seguridad informática.

Referencias

[1] Spafford E., "The Internet Worm Program: An Analysis". Purdue University. (1988).

[2] Young A., Yung M., "Malicious Cryptography Exposing Cryptovirology". Wiley Publishing, Inc. (2004).

[3] Moore C., "Detecting Ransomware with Honeypot techniques". Cybersecurity and Cyberforensics Conference, IEEE. (2016).

[4] Savage K., Coogan P., "The evolution of ransomware". Symantec. (2015).

[5] Mehmood S., "Enterprise Survival Guide for Ransomware Attacks". The SANS Institute. (2016).

[6] Richardson R., North N., "Ransomware: Evolution, Mitigation and Prevention". International Management Review. Vol. 13. No. 1. (2017).
http://scholarspress.us/journals/IMR/pdf/IMR-1-2017.%20pdf/IMR-v13n1art2.pdf

[7] Hampton N., Baig Z., "Ransomware: Emergence of the cyber-extortion menace". Edith Cowan University Australian Information Security Management Conference. (2015).
http://ro.ecu.edu.au/cgi/viewcontent.cgi?article=1179&context=ism

[8] Lupu E., Sgandurra D., Muñoz-González M., Mohsen R., "Automated Dynamic Analysis of Ransomware: Benefits, Limitations and use for Detection". arXiv.org E-print Archive arXiv:1609.03020 [cs.CR]. (2016).

[9] MalwareLabs, "Petya – Taking Ransomware To The Low Leve". (2017).

[10] Scott J., Spaniel D., "Cerber & KeRanger: The Latest Examples of Weaponized Encryption". Institute for Critical Infrastructure Technology (ICIT). (2016).

[11] Proofpoint Staff, "Philadelphia Ransomware Brings Customization to Commodity Malware". (abril 2017).

[12] Fact Sheet: ETERNALBLUE exploit and DOUBLEPULSAR backdoor. Australian Cyber Security Centre. (mayo 2017).

[13] Sanggeun S., Bongjoon K., Sangjun L., "The effective Ransomware Prevention Technique Using Process Monitoring on Android Platform. Soongsil University. (2016).

[14] Microsoft. “Server Message Block Overview”. Documentación official: https://docs.microsoft.com/es-es/previous-versions/windows/server/hh831795(v=ws.11)

Criptografía Ligera e Internet de las Cosas

Mg. Jorge Eterovic; Esp. Marcelo Cipriano; Lic. Edith García.

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; cipriano1.618; edithxgarcia}@gmail.com

RESUMEN

Es notorio el incremento de los usos y aplicaciones de la llamada Internet de las Cosas (IoT por sus siglas en inglés). La razón de ser de esta plétora de tecnologías concurrentes permite la conexión de objetos de distinta naturaleza e índole, a través de Internet.

La incidencia de su uso sobre la humanidad se vislumbra como un cambio de paradigma[1] al que ya la sociedad se está acostumbrando: no se sorprende ante la obtención de datos médicos en tiempo real por medio de dispositivos e-Health[1], el implante y uso de chips subcutáneos para la identificación y rastreo de personas, las zapatillas con GPS o las heladeras que hacen el pedido al supermercado de la mercadería faltante[2], entre otros aparatos. La tecnología IoT va penetrando en el entorno humano y todo hace pensar que esto recién es el principio.

Sin embargo de la mayoría de las personas ignora que por la propia naturaleza de estos dispositivos -que deben llevar adelante su tarea en contextos reducidos en tamaño, potencia de cómputo, consumo eléctrico, tamaño de las baterías, cantidad de memoria asignada, entre otras limitaciones- se detecta una reducción de la seguridad y privacidad de las comunicaciones y datos que se transmiten, procesan y almacenan. Incluso existen dispositivos que carecen completamente de ellas.

Es por ello que este proyecto propone llevar adelante el estudio de algoritmos pertenecientes a la llamada Criptografía Liviana [3], los que por su diseño pueden ser ejecutados en este tipo de dispositivos.

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 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 (Código VRID 1935 – Código Académico 100091) con una duración de 2 años (2019-2020).

[1] E-Salud: cuidados sanitarios apoyados en dispositivos TIC´s como pueden ser marcapasos, bombas de insulina, implantes cocleares, etc.

1. INTRODUCCIÓN

Aunque la llamada Internet de las Cosas aparece ante la sociedad como recién llegada al mundo, este concepto no es nuevo. El primer dispositivo conectado a internet fue una máquina expendedora de Coca Cola en la Universidad Carnegie Mellon, en el año 1982. Inocentemente comenzó como una manera de saber cuántas latas de producto aún tenía la máquina, si ya estaban frías o había que esperar para ir a comprarlas. Fue recién en 1999 que el término "Internet de las Cosas" fue propuesto por Kevin Ashton – que trabajaba en el Auto-ID Center del MIT pues allí le llevaban adelante investigaciones en los campos de nuevas tecnologías inalámbricas llamadas RFID[2] y WSN[3]- en una conferencia dictada en la empresa Procter & Gamble en 1999.

Enormes cantidades de información son recolectadas, procesadas y transmitidas, por estos dispositivos, a través de la Internet. Cabe preguntar acerca de los mecanismos de seguridad que tales equipos poseen. Dadas las limitaciones intrínsecas que estos aparatos tienen, existe la posibilidad que no se cuenten con los mecanismos de seguridad adecuados.

La sociedad debiera llevar adelante un extenso análisis en torno a la privacidad y seguridad de la información que se manipula en los entornos IoT[4], fuera de su control y al cuidado de empresas, gobiernos propios y terceros. Es lícito preguntar ¿qué hacen las empresas y organismos con la información que recolectan? ¿Cuál es el límite a partir del cual comienza a violarse la privacidad?

Los asistentes virtuales Siri, Alexia o diferentes Smarts TV son excelentes herramientas para personas con movilidad reducida. Nada haría presuponer en ello una amenaza, todo lo contrario. Hasta que alguien se percató que esos dispositivos inteligentes se mantienen en estado de escucha permanente a la espera de ser activados por la voz de sus propietarios. De esa manera fueron blanco de ataques, tal como quedó en evidencia en la presentación dada por Aaron Grattafiori y Josh Yavor en la conferencia Black Hat de 2013 en la que se mostró la manera de acceder al micrófono o la cámara incorporados a un smart tv para realizar espionaje o voyeurismo en el seno de un hogar cualquiera.

Frente a todo lo expuesto, pocos son los mecanismos viables para dotar de confidencialidad a las comunicaciones. Uno de ellos es el uso de algoritmos de cifrado que por su robustez y confiabilidad son una de las mejores alternativas.

En el contexto de la Criptografía Liviana[5] se pueden encontrar algoritmos de clave pública y clave privada, Block Ciphers[6-8] y Stream Ciphers[9-12] como así también algoritmos para la Gestión de Claves, Firma Digital y funciones Hash[13-15]. Es decir todo el conjunto de funcionalidades que la Criptografía convencional ofrece, pero capaces de correr en contextos reducidos como lo son los dispositivos IoT y sin que ello menoscabe su robustez y performance.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

La metodología a implementar es principalmente de tipo analítico y podría tener un correlato secundario de tipo experimental. Al aspecto analítico le corresponde el estudio de los algoritmos criptográficos, sus aspectos y propiedades matemáticas, como así también la propuesta de mejoras en la implementación de los mismos, aplicados a Internet de las Cosas.

En cuanto al aspecto experimental, se podrán llevar adelante pesquisas de laboratorio al aplicar el cifrado correspondiente a un determinado algoritmo y la observación de su comportamiento en ese entorno reducido de trabajo. Para tal fin podrían llevarse adelante mecanismos de virtualización y/o simulación de los equipos y ecosistema de software donde el algoritmo desempeñará sus funciones. Que podrían ser motivo de futuras líneas de investigación y trabajos netamente experimentales.

[2]RFID: Identificación por Radio Frecuencia (Radio Frequency IDentification).

[3] WSN: Wireless sensor network.

Asimismo cabe aclarar que los algoritmos criptográficos se encontrarán compartiendo hardware y software con otros algoritmos, constituyendo un "ecosistema" muy limitado y reducido, donde el comportamiento de los mismos podría recibir influencias negativas no previstas por sus creadores. Es por ello que podrían percibirse diferencias en el comportamiento esperado acerca de la velocidad y el desempeño, al ejecutarse en dispositivos de este tipo. Y es por ello que tal vez se requiera de la virtualización y/o simulación de los entornos y equipos donde el algoritmo podría desempeñarse.

Para llevar adelante el proyecto se proponen varias líneas de acción para conducir la investigación y el desarrollo del mismo:

- búsqueda de información y relevamiento de novedades acaecidas en el mundo de la Criptografía Liviana y su relación a Internet de las Cosas, dado que es un campo de investigación en constante ebullición, con la aparición de novedades a ritmo vertiginoso.
- análisis de los Protocolos de Comunicaciones en IoT que utilicen criptografía y ofrezcan algún tipo de seguridad.
- Relevamiento de los algoritmos criptográficos livianos más relevantes, tipo, modos de uso, longitud de la clave, mecanismos de cifrado, aplicaciones y usos.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Este proyecto persigue como objetivo profundizar el estudio y análisis de algoritmos criptográficos pertenecientes a la llamada Criptografía Liviana que puedan ser usados en dispositivos de la Internet de las Cosas y que permiten asegurar la información, tanto en su transmisión como almacenamiento.

Se tratará de sugerir mejoras en los mismos, de manera que el comportamiento de la seguridad por medio de la criptografía sea más performante, sugiriendo mejoras en cuanto a la velocidad y robustez en las comunicaciones, lo que podrá conllevar como consecuencia una reducción en el consumo de la batería, espacio en memoria y demás indicadores de la Criptografía Ligera.

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 este proyecto que recién inicia se incorporarán próximamente una docente investigadora de larga experiencia y 2 alumnos que se encuentran promediando la carrera de Ingeniería en Informática. Esto redundará en un aumento del activo académico e investigativo representado por su cuerpo de docentes investigadores, como así también sembrando las bases para la investigación del futuro, a través de la participación de alumnos de la Facultad de Ingeniería.

5. BIBLIOGRAFÍA

[1] 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.

[2] http://tn.com.ar/tecno/f5/ces-2016-las-heladeras-del-futuro-conectadas-y-con-multiples-sensores_647274. Consultada el 28/2/19.

[3] ISO/IEC 29192. Information technogy - Security techniques - Lightweight Cryptography. 2012. https://www.iso.org.

[4] Puente García, M. "Iniciativas y mejores prácticas de seguridad para el IoT". Instituto Nacional de Ciberseguridad Español (INCIBE). 2017. https://www.incibe-cert.es/blog/ iniciativas-y-mejores-practicas-seguridad-el-iot. Consultada el 28/2/19.

[5] Panasenko, S.; Smagin, S. "Lightweight Cryptography: Underlying Principles and Approaches". International Journal of Computer Theory and Engineering, Vol. 3, No. 4, August 2011.

[6] Satoh, A.; Morioka, S. "Hardware-Focused Performance Comparison for the Standard Block Ciphers AES, Camellia, and

Triple-DES". Conference: Information Security, 6th International Conference, ISC 2003, Bristol, UK, October 1-3, 2003, Proceedings.
[7] Beaulieu, R.; Shors, R.; Smith, J.; Treatman-Clark, S.; WeeksWeeks, B.; Wingers. L. "The SIMON and SPECK Families of Lightweight Block Ciphers." Cryptology EPrint Archive. International Association for Cryptologic Research, 19 June 2013.
[8] Dworkin, M. "NIST SP 800-38B, Recommendation for Block Cipher Modes of Operation: The CMAC Mode for Authentication." NIST Computer Security Resource Center. National Institute of Standards and Technology, Spring (2005).
[9] Daniel J. Bernstein. "The Salsa20 family of stream ciphers" . URL:http://cr.yp.to/ papers.html# salsafamily. (2007).
[10] Babbage, S.; Dodd, M. "The MICKEY stream ciphers". In New Stream Cipher Designs. Pp. 191-209. Springer Berlin Heidelberg. (2008).
[11] Hell, M.; Johansson, T.; Meier, W. "Grain: a stream cipher for constrained environments". International Journal of Wireless and Mobile Computing, 2, pp. 86-93 (2007).2
[12] De Canniere, C.; Preneel, B. "Trivium. New Stream Cipher Designs (pp. 244-266). Springer Berlin Heidelberg. (2008).
[13] Kavun, E. B., & Yalcin, T. "On the suitability of SHA-3 finalists for lightweight applications". The Third SHA-3 Candidate Conference. (2012).
[14] Hirose, S., Ideguchi, K., Kuwakado, H., Owada, T., Preneel, B., & Yoshida, H. "A lightweight 256-bit hash function for hardware and low-end devices: Lesamnta-LW". International Conference on Information Security and Cryptology. Pp. 151-168. Springer Berlin Heidelberg (2010).
[15] Guo, J.; Peyrin, T.; Poschmann, A. "The PHOTON family of lightweight hash functions". Advances in Cryptology–CRYPTO 2011 (pp. 222-239). Springer Berlin Heidelberg (2011).

CRIPTOGRAFÍA POST-CUÁNTICA INTEGRADA EN SSL/TLS Y HTTPS

Diego Cordoba [1], Miguel Méndez-Garabetti[2]

[2] Univerisdad de Mendoza, Dirección de Posgrado, Facultad de Ingeniería
diego.cordoba@um.edu.ar, miguel.mendez@um.edu.ar

[1]Universidad de Mendoza, Facultad de Ingeniería, Subsede San Rafael

RESUMEN

La adopción del protocolo HTTPS para brindar seguridad a gran parte del tráfico web de Internet pone el foco en los algoritmos de cifrado que proveen las implementaciones de SSL/TLS.

La capacidad de cómputo de los ordenadores actuales permite concluir que algoritmos asimétricos como RSA son seguros. No obstante, estos mecanismos son vulnerables al criptoanálisis cuántico, por lo que disponer de una computadora cuántica pondría en peligro la seguridad de los datos cifrados con algoritmos asimétricos como RSA.

La computación cuántica es un hecho, y que es cuestión de tiempo para que se fabriquen computadoras cuánticas de cientos de qubits, por lo que es necesario encontrar alternativas a los algoritmos de cifrado y autenticación clásicos. La criptografía postcuántica cubre esta necesidad proveyendo una serie de algoritmos resistentes a ataques cuánticos. Existen proyectos que aportan librerías de desarrollo bajo licencias de código abierto. OQS provee una librería de algoritmos resistentes, y su integración en protocolos y aplicaciones como OpenSSL, una de las implementaciones de SSL/TLS de código abierto más utilizadas.

El presente trabajo de investigación analiza la integración de estas implementaciones de OpenSSL en el servidor web Apache2 para brindar un servicio HTTPS resistente a ataques cuánticos.

Palabras clave: criptografía; post cuántico; seguridad informática; criptografía asimétrica

CONTEXTO

El presente trabajo de I+D se desarrolla como proyecto se tesis de posgrado de la Maestría en Teleinformática, Dirección de Posgrado, perteneciente a la Facultad de Ingeniería de la Universidad de Mendoza, (Ciudad, Mendoza). El presente proyecto fue presentado como propuesta de tesis.

1. INTRODUCCIÓN

En la actualidad las comunicaciones en Internet, particularmente en el tráfico web, tienden a ser cifradas para mantener la confidencialidad de los datos del usuario, y la autenticación e integridad de los mismos entre los nodos que conforman el enlace. El incremento de adopción del protocolo HTTPS (HTTP over SSL/TLS) [1] marca tendencia, y algunos navegadores como Chrome o Firefox comienzan a advertir cuando los sitios accedidos no son seguros [2] [3] e incluso los buscadores en Internet comienzan a tener en cuenta el uso de HTTPS en el posicionamiento [4].

SSL/TLS le brindan a HTTP algoritmos de autenticación basados en certificados digitales X.509 y criptografía asimétrica. La criptografía asimétrica se basa en algoritmos matemáticos fundamentados en supuestos de complejidad computacional, es decir, problemas matemáticos muy simples de calcular, pero extremadamente difícil de revertir. De esta manera la generación de claves asimétricas, por ejemplo, pueden llevarse a cabo con operaciones denominadas "de una sola vía", tales como la multiplicación de dos números primos grandes. El producto de estos dos números primos es relativamente sencillo de computar para un procesador actual, pero encontrar estos dos números primos que dieron origen al producto resulta sumamente difícil aplicando algoritmos de fuerza bruta con las capacidades de cálculo actuales. Si a esto se agrega que las claves de cifrado y autenticación asimétrica suelen cambiar periódicamente dependiendo de la implementación de SSL/TLS que se utilice, y/o su configuración, podría decirse que romper un cifrado asimétrico de clave grande resulta, al día de hoy, prácticamente imposible para un atacante que necesita conseguir la clave de cifrado en un tiempo limitado.

Por otro lado, una computadora cuántica podría encontrar la clave de cifrado de manera relativamente rápida, y vulnerar, de esta forma, cualquier tráfico cifrado y autenticado con mecanismos basados en criptografía asimétrica.
El primer algoritmo cuántico no trivial que demostró un potencial de crecimiento exponencial de velocidad sobre los algoritmos clásicos fue el algoritmo de Shor [5]. Este algoritmo permite descifrar un mensaje encriptado mediante RSA [6] descomponiendo en factores la clave pública, que es producto de dos números primos grandes, en un tiempo $O((\log N)^3)$, siendo N el numero primo que representa la clave pública. Por su parte, los algoritmos clásicos no pueden realizar dicha factorización en un tiempo menor a $O((\log N)k)$ para ningún k, por lo que RSA sigue siendo considerado, en la actualidad, un algoritmo seguro [7] [8].
El algoritmo de Shor, al ser un algoritmo cuántico, permite obtener el resultado de manera probabilística y con un determinado grado de acierto de acuerdo a la cantidad de iteraciones a la que se lo someta. En la práctica Isaac Chuang en diciembre de 2001, liderando un grupo de trabajo de computación cuántica de IBM, logró factorizar el número 15 mediante una computadora cuántica de 7 qubits, y en marzo de 2016 un grupo de investigadores del MIT, entre los que se encontraba el mismo Chuang, también pudo factorizar el numero 15 con un 99% de certeza en una computadora cuántica de 5 qubits.

La aparición de las primeras computadoras cuánticas de cientos de qubits dejará obsoletos a la mayoría de los algoritmos asimétricos actuales, tales como RSA, DSA o ECDSA [7] [9] [10] [11].

Afortunadamente existen algoritmos asimétricos que hacen uso de mecanismos que son invulnerables al criptoanálisis cuántico. Estos algoritmos definen la criptografía post-cuántica, o criptografía resistente a ataques cuánticos, son considerados seguros y a ninguno de ellos se le ha podido aplicar el algoritmo de Shor.

Los algoritmos post-cuánticos pueden clasificarse en [7]:

1. *Criptografía basada en hash*, que incluye sistemas criptográficos como las firmas de Lamport y el esquema de firmas de Merkle.

2. *Criptografía basada en código*, que incluye el esquema de firmas de McEliece y códigos aleatorios Goppa [12].

3. *Criptografía basada en sistemas de ecuaciones multivariable*, que incluye el esquema Rainbow [13].

4. *Criptografía basada en enrejado*, que incluye algoritmos de intercambio de claves de aprendizaje con errores, NTRU [14] y BLISS [15].

5. *Criptografía simétrica basada en clave secreta de Rijndael*, más conocido como AES (Advanced Encryption Standard).

Para poder llevar a la práctica estos algoritmos es necesario que las implementaciones de protocolos SSL/TLS los soporte. Uno de los proyectos que se mantiene más activo en este aspecto es OQS – Open Quantum Safe [16].

El objetivo de OQS es dar soporte al desarrollo y *prototipado* de implementaciones de algoritmos resistentes a ataques cuánticos. Para ello desarrollaron liboqs, una biblioteca de algoritmos post-cuánticos en lenguaje C, y que sirve de API para utilizar dichos algoritmos, y está liberada bajo licencia de código abierto del MIT. OQS mantiene dos ramas de liboqs, la rama master, y la rama NIST. La rama master está enfocada en algoritmos seleccionados de encapsulamiento de claves y firma digital, y tiene determinados criterios de aceptación de las implementaciones. Por su parte, la rama NIST se centra en incorporar algoritmos remitidos al proyecto de estandarización de criptografía postcuántica del NIST [17]. Además, el proyecto OQS provee encapsulamientos / wrappers para lenguajes específicos como C#, C++ y Python.

Además de desarrollar y mantener liboqs, el proyecto OQS también realiza tareas de integración de su biblioteca en algunos protocolos e implementaciones de código abierto que pasan a beneficiarse con la incorporación de algoritmos postcuántico. Tal es el caso de OpenSSL, una de las implementaciones de SSL/TLS de código abierto más utilizadas actualmente, otorgándole al desarrollo los mecanismos resistentes a ataques cuánticos para la autenticación y el intercambio de claves.

Ademas, la biblioteca liboqs ha sido utilizada por proyectos de terceros, como ser la VPN postcuántica experimental de Microsoft, basada en OpenVPN [18], o el cliente beta de VPN de Mullvad [19], que utiliza intercambio de claves postcuántico.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

Los algoritmos de cifrado, autenticación e intercambio de claves postcuánticos tienen una gran cantidad de aplicaciones en diferentes herramientas que hagan uso de criptografía asimétrica clásica.
Particularmente en este trabajo se lleva a cabo un análisis de la bibliografía existente, y las principales implementaciones de algoritmos resistentes a ataques cuánticos, y se centra el foco de atención el la librería liboqs del proyecto OQS, y las implementaciones de OpenSSL provistas por el proyecto que integran esta librería.
Los aspectos que se pretenden desarrollar con la presente investigación incluyen:

1. Analizar y documentar los algoritmos de cifrado, autenticación e intercambio de claves resistentes a ataques cuánticos disponibles en la actualidad.
2. Estudiar y realizar pruebas de concepto con las librerías de cifrado que soporten algoritmos postcuánticos, específicamente la librería liboqs del proyecto OQS, tanto en la rama master como en la rama NIST.
3. Compilar la suite OQS-OpenSSL tanto en la versión 1.1.1-stable como en la versión 1.0.2-stable (ambas liberadas en Noviembre de 2018) utilizando las bibliotecas liboqs de la rama master y nist, y realizar pruebas de concepto y conectividad cliente-servidor local.
4. Realizar pruebas de conectividad y negociación TLS de los algoritmos postcuánticos o híbridos para intercambio de claves provistos por la versión OQS-OpenSSL v1.1.1-stable (para TLS v1.3) y v1.0.2-stable (para TLS1.2).
5. Realizar la compilación de la última versión del servidor web Apache2 httpd (v2.4.38) junto con su módulo SSL/TLS, y realizar pruebas de concepto de conectividad, renegociación de *cipher suites* e intercambio de claves para ambas versiones de OQS-OpenSSL.
6. Realizar pruebas de carga tanto en las implementaciones de OQS-OpenSSL como en el servicio HTTPS de Apache2, respecto del establecimiento de conexiones cifradas y el intercambio de claves, y medir el rendimiento de los algoritmos postcuánticos comparado con el rendimiento de los algoritmos tradicionales provistos por OpenSSL.

3. RESULTADOS OBTENIDOS/ ESPERADOS

Al momento de publicar este artículo se han llevado a cabo satisfactoriamente las siguientes pruebas:

1. Compilación, configuración y puesta en marcha de OQS-OpenSSL-1.0.2-stable con liboqs rama nist.
2. Compilación, configuración y puesta en marcha de OQS-OpenSSL-1.1.1-stable con liboqs rama master.
3. Ejecución exitosa del comando de prueba de rendimiento *speed* de la suite OpenSSL, tanto para la v1.1.1-stable como para la v1.0.2-stable provistas por OQS.
4. Generación de claves asimétricas y certificados digitales X.509 haciendo uso de algoritmos asimétricos resistentes a ataques cuánticos.
5. Compilación, configuración y puesta en marcha del servidor web Apache2 con el módulo SSL/TLS compilado en base a la implementación OQS-OpenSSL v1.0.2-stable provista por el proyecto OQS.

Las pruebas realizadas sobre la versión v2.4.38 de Apache combinada con OQS-OpenSSL v1.0.2-stable generan inconsistencias en la detección de algoritmos postcuánticos de autenticación e intercambio de claves, negociación y renegociación de algoritmos durante el establecimiento del canal seguro de TLS v1.2. Esta inconsistencia queda pendiente de depuración, análisis y posible solución.

Por su parte, al finalizar el proyecto también se pretende obtener conclusiones acerca del rendimiento de los algoritmos postcuánticos de generación de claves comparados con los algoritmos clásicos partiendo de la información obtenida con el comando *speed* de OpenSSL y de herramientas de pruebas de stress específicas.

4. FORMACIÓN DE RECURSOS HUMANOS

La línea de I+D presentada está vinculada con el desarrollo de una tesis de posgrado por parte del Ing. Diego Córdoba, quien es estudiante de la Maestría en Teleinformática de la Universidad de Mendoza.

5. BIBLIOGRAFÍA

[1] EC Council, *Network Defense: Fundamentals and Protocols*. EC Council Press, 2010.

[2] C. Cimpanu, "Firefox Prepares to Mark All HTTP Sites 'Not Secure' After HTTPS Adoption Rises," 18-Dec-2017. [Online]. Available: https://www.bleepingcomputer.com/news/software/firefox-prepares-to-mark-all-http-sites-not-secure-after-https-adoption-rises/. [Accessed: 02-Aug-2019].

[3] Chromium Blog - Google, "A secure web is here to stay," 02-Aug-2018. [Online]. Available: https://blog.chromium.org/2018/02/a-secure-web-is-here-to-stay.html. [Accessed: 08-Feb-2019].

[4] Z. A. Bahajji and G. Illyes, "HTTPS as a ranking signal," 2014. [Online]. Available: https://webmasters.googleblog.com/2014/08/https-as-ranking-signal.html.

[5] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *Journal SIAM Journal on Computing*, vol. 26, no. 5, pp. 1484–1509, 1997.

[6] R. L. Rivest, A. Shamir, and L. Adlemar, *A Method for Obtaining Digital Signatures and Public-Key Cryptosystems*. EEUU: Communications of the ACM, 1987.

[7] D. J. Bernstein, J. Buchmann, and E. Dahmen, *Post Quantum Cryptography*, 1ra ed. Berlin: Springer, 2009.

[8] E. W. Weisstein, "RSA-640 Factored," *MathWorld Headline News*, 05-Nov-2005.

[9] (European Telecommunications Standards Institute) ETSI, *Quantum Safe Cryptography and Security*, ETSI White Paper No. 8. France: ETSI, 2015.

[10] T. Takagi, Ed., *Post-Quantum Cryptography*, PQCrypto 2016. Fukuoka, Japón: Springer, 2016.

[11] R. A. Perlner and D. A. Cooper, *Quantum Resistant Public Key Cryptography: A Survey*. Maryland, EEUU: National Institute of Standards and Technology, 2009.

[12] D. J. Bernstein and T. Lange, "Attacking and Defending the McEliece cryptosystem.," presented at the PQCrypto 2008, 2008.

[13] P. Oechsli, *Making a faster cryptanalytical Time-Memory Trade-of. Advances in cryptology: Proceedings of CRYPT*. 2003.

[14] OnBoard Security, "NTRU Post Quantum Cryptography," *NTRU Post Quantum Cryptography*, 2018. [Online]. Available: https://www.onboardsecurity.com/products/ntru-crypto.

[15] L. Ducas, A. Durmus, T. Lepoint, and V. Lyubashevsky, "BLISS: Bimodal Lattice Signature Schemes," presented at the CRYPTO 2013, 2013.

[16] D. Stebila and M. Mosca, *Post-Quantum Key Exchange for the Internet and the Open Quantum Safe Project*. EEUU: Department of Computing and Software, Mc Master University, 2017.

[17] NIST (National Institute of Standards and Technology), "Post-Quantum Cryptography Project," 03-Jan-2017. [Online]. Available: https://csrc.nist.gov/Projects/Post-Quantum-Cryptography/Post-Quantum-Cryptography-Standardization. [Accessed: 13-Mar-2019].

[18] Microsoft, "Welcome to the PQCrypto-VPN project!," *Welcome to the PQCrypto-VPN project!*, Jul-2018. [Online]. Available: https://github.com/Microsoft/PQCrypto-VPN.

[19] Amagicom AB, "Introducing a post-quantum VPN, Mullvad's strategy for a future problem," 08-Dec-2017. [Online]. Available: https://www.mullvad.net/en/blog/2017/12/8/introducing-post-quantum-vpn-mullvads-strategy-future-problem/. [Accessed: 12-Mar-2019].

Criptografía y Sistemas Móviles

Castro Lechtaler, Antonio[1, 2]; Cipriano, Marcelo[1;3]; García, Edith[1],
Liporace, Julio[1]; Maiorano, Ariel[1]; Malvacio, Eduardo[1]; Tapia, Néstor[1];

[1]Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Escuela Superior Técnica, Facultad del Ejército. Universidad de la Defensa Nacional UNDEF.
[2] CISTIC/FCE - Universidad de Buenos Aires.
[3] Departamento 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

Este proyecto persigue el estudio, diseño y desarrollo de un Algoritmo de Cifrado en Cadena o Stream Cipher, que pueda ser ejecutado sobre un sistema móvil (teléfono celular, tablet, radios que operan en bandas VHS, UHF o cualquier otro dispositivo de similares características) para así poder ofrecer confidencialidad a la información transmitida.
Tal algoritmo deberá ofrecer capacidades especiales, como trabajar en altas velocidades de cifrado/descifrado, robustez y resistencia a los ataques criptoanalíticos conocidos, tales como el Criptoanálisis Diferencial [1], Criptoanálisis Lineal [2], Cube Attack Cryptanalysis [3-4] y otros [5-6].
Para ello se seguirá la filosofía de diseño y construcción en la que atendiendo los ataques conocidos el algoritmo se diseña resistente a dichos ataques [7].
También se tendrá en cuenta que dicho algoritmo debe ejecutarse en forma eficiente en dispositivos reducidos en tamaño, potencia de cómputo, capacidad de memoria y consumo eléctrico, entre otras. Ya que muchos de los sistemas móviles priorizan la movilidad, liviandad de los equipos, reducidos consumos eléctricos para prolongar la carga de la batería y demás limitaciones que conspiran contra el normal desempeño de un algoritmo criptográfico.

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 Facultad de Ingeniería del Ejército (FIE) "Gral. Div. Manuel N. Savio", 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

A la población en general parece no importarle (o al menos eso parece) la confidencialidad de sus comunicaciones. La información que nuestros dispositivos de redes y comunicaciones procesan y transmiten, están sometidos riesgos más allá de los ataques de hackers u otros. El riesgo que gobiernos y empresas puedan acceder fácilmente a la información. Esto quedó expuesto a nivel mundial en la masacre de San Bernardino de 2015[1]. Más allá del lamentable suceso, se originó una discu-

[1] Hecho luctuoso que se realizó durante un banquete del Departamento de Salud Pública realizado en un centro para personas con capacidades diferentes del condado de San Bernardino, California (Estados Unidos) el 2 de Diciembre de 2015. Allí 2 terroristas dispararon contra los comensales hiriendo al menos a 21 personas y asesinando a 14. Uno de ellos era empleado en ese lugar y compañero de muchas de las víctimas.

sión ética y legal entre la Oficina Federal de Investigación[2] y la empresa Apple pues la policía encontró un teléfono iPhone 5C, propiedad de uno de los terroristas. El FBI le exigía a la empresa que "abra" el móvil que se encontraba bloqueado y su contenido cifrado con AES-256[3]. En particular solicitaban que se les entregue una especie de "llave maestra" del cifrado, una "puerta trasera" del sistema operativo iOS del teléfono o en su defecto que la empresa Apple intente hackear su propio teléfono, para acceder a la información almacenada en él.

La empresa se negaba a hacerlo alegando que no podía atentar contra la confidencialidad de sus usuarios y que además eso iría contra los intereses comerciales de la misma ya que entonces sus clientes podrían dejar de usar sus teléfonos o dejar de comprarlos.

La controversia pronto escaló hasta la justicia estadounidense. Luego de un tiempo litigando en los tribunales, discutiendo en las redes sociales y los medios de comunicación, el gobierno desistió de la demanda: ya lo habían resuelto por su cuenta… es decir… pudo acceder a la información almacenada en el equipo.

Como usuarios tenemos pocas herramientas a nuestro favor que nos permitan proteger nuestra información. Desde la empresa que diseña, fabrica y comercializa los equipos (tanto computadoras, tablets, móviles y demás) hasta la infraestructura de Internet, nuestra información es procesada, almacenada y transmitida con escaso o nulo control sobre ella.

El Grupo de Investigación en Criptografía y Seguridad Teleinformática (GICSI) por medio de este proyecto intenta ofrecer un mecanismo de seguridad para preservar la confidencialidad de la información mediante el uso de un mecanismo de cifrado robusto y adaptado a sistemas móviles.

Muchos de estos sistemas tienen amplios recursos de hardware y software (algunos teléfonos móviles sorprenden por las altísimas prestaciones que ofrecen, gracias a la potencia de su hardware), otros en cambio, tienen reducidos recursos para lograr sus objetivos. Tal es el caso de sistemas que al priorizar la movilidad o portabilidad de los mismos, deben reducir su tamaño, peso, capacidad de almacenamiento, cómputo y consumo energético, por mencionar algunas de las restricciones.

Esa reducción no debiera influir en la disminución de la confidencialidad que tales equipos requieran. Existen mecanismos criptográficos que permiten ofrecer la robustez necesaria aún en contextos limitados.

CriptoLab desde el ámbito académico ha realizado algunas propuestas para vehículos aéreos no tripulados del Proyecto LIPAM del Ejército Argentino, los cascos de realidad aumentada del Proyecto[4] RAIOM del Centro de Investigaciones para la Defensa - CITEDEF[5]. También se pueden mencionar otros sistemas y vehículos militares, como el PANHARD francés que el Ejército y otras fuerzas poseen y que la FIE recibió como encargo, llevar adelante su modernización.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Para llevar adelante el proyecto, se ha dividido en 4 etapas de investigación y desarrollo:

- Estudio y análisis de algoritmos que satisfacen los requerimientos y condiciones de entorno del proyecto.
- Personalización, diseño, desarrollo del algoritmo e inclusión de mejoras.
- Determinación de las propiedades criptológicas.
- Ejecución de los tests y demás pruebas.

[2] En inglés: Federal Bureau of Investigation. FBI.

[3] AES: en inglés Advanced Encryption Standard. Algoritmo criptográfico de 256 bits de clave privada, que se convirtió en un estándar del NIST en 2002. A la actualidad es considerado seguro y resistente a los ataques conocidos… ¿o ya no?

[4] RAIOM: Realidad Aumentada para la Identificación de Objetivos Militares.

[5] CITEDEF: El Instituto de Investigaciones Científicas y Técnicas para la Defensa; ex Instituto de Investigaciones Científicas y Técnicas de las Fuerzas Armadas (CITEFA)

3. RESULTADOS OBTENIDOS/ESPERADOS

La persigue el diseño de un mecanismo de cifrado tipo Stream Cipher que dote de confidencialidad a las comunicaciones de un Sistema Móvil. Que pueda además demostrar su fortaleza al resistir los ataques criptoanalíticos conocidos.

La realización de un desarrollo propio y nacional no sólo permitirá ofrecer una capa más de protección a la información respecto a la confidencialidad de la misma por medio de un algoritmo criptográfico, sino que además permitirá ahorrar recursos económicos. El costo de desarrollar un algoritmo puede apreciarse más beneficioso que la adquisición de uno extranjero, sobre todo por el alto costo de los mismos, valuados en moneda foránea.

4. FORMACIÓN DE RECURSOS HUMANOS

Los docentes investigadores 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. Es por ello que 3 de ellos han demostrado su interés y se han sumado en calidad de colaboradores. En particular, el alumno Leiras, Facundo ha presentado su postulación en 2018 para la beca "Estímulo a las Vocaciones Científicas" (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones requeridas[8]. La misma le ha sido otorgada, iniciando en breve sus actividades respectivas.

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 y de la institución en la cual desarrollan sus actividades científico-docentes.

Por último y 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] 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.

[2] 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.

[3] Dinur, Itai; Shamir, Adi (2009-01-26). "Cube Attacks on Tweakable Black Box Polynomials" (PDF). Cryptology ePrint Archive. ePrint 20090126:174453.

[4] 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.

[5] Pasalic, E.; *On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers*; IEEE Transactions on Information Theory. Vol. 55 Ed.7º, 2009.

[6] 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.

[7] 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.

[8] http://evc.cin.edu.ar/informacion consultada el 23/2/2018.

Detección de vulnerabilidades en especificaciones de contratos inteligentes de la plataforma Ethereum

Mauro C. Argañaraz(1), Mario M. Berón(1), Miguel A. Bustos(1), Pedro Rangel Henriques(2) &Daniel Riesco(1)
(1)Departamento de Informática - Facultad de Ciencias Física Matemáticas y Naturales (FCFMyN) – Universidad Nacional de San Luis
(2)Universidade do Minho - Braga, Portugal
marganaraz@gmail.com(1), {mberon, mabustos, driesco}@unsl.edu.ar(1), pedrorangelhenriques@gmail.com(2)

RESUMEN

Ethereum es el principal ecosistema basado en blockchain que proporciona un entorno para codificar y ejecutar contratos inteligentes, propiedad que en estos últimos años recibió gran atención por las nuevas aplicaciones comerciales y, también, por la comunidad científica. El proceso de escritura de contratos seguros y de buen desempeño en la plataforma Ethereum es un gran desafío para los desarrolladores. Implica la aplicación de paradigmas de programación no convencionales debido a las características inherentes de la ejecución de programas de computación distribuida. Además, los errores en los contratos desplegados pueden tener graves consecuencias debido al acoplamiento inmediato del código del contrato y las transacciones financieras. El manejo directo de los activos significa que las fallas tienen más probabilidades de ser relevantes para la seguridad y tienen mayores consecuencias económicas que los errores en las aplicaciones típicas. Una serie de reportes de incidentes muestran que los problemas de seguridad se han utilizado con propósitos fraudulentos.

En este artículo, se describe una línea de investigación que se enfoca en fortalecer los aspectos de seguridad, basado en una base sólida de diseño, patrones de código establecidos y probados que faciliten el proceso de escritura de código funcional y libre de errores, para proporcionar una herramienta que permita analizar y detectar falencias de manera automática.

Palabras Claves: blockchain, Ethereum, smart contract, Solidity, static analysis tool, verification, security patterns.

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

El año 2017 se caracterizó por la popularidad mundial que alcanzó Bitcoin [1] llegando a su valor de mercado más alto en su breve recorrido desde su creación, hito que coincidió con el reconocimiento de Wall Street al incluirlo en los mercados de futuros. Ante la merma en el precio durante 2018, la cadena de bloques (blockchain), tecnología que subyace a Bitcoin y al resto de las criptomonedas, llegó a la cresta de la ola IT. Sin embargo, de ahora en más, el

foco del mercado y de la comunidad científica inevitablemente se concentrará en los contratos inteligentes, ya que permiten que partes no confiables manifiesten términos de contrato en el código del programa y, por lo tanto, eliminan la necesidad de un tercero confiable.

Blockchain es una tecnología que se basa en una combinación de criptografía, redes y mecanismos de incentivos para respaldar la verificación, ejecución y registro de transacciones entre diferentes partes. En términos simples, las plataformas de blockchain se pueden ver como bases de datos descentralizadas que ofrecen propiedades muy atractivas. Esto incluye la inmutabilidad de las transacciones almacenadas y la creación de confianza entre los participantes sin un tercero. Esto hace que esta arquitectura sea adecuada como un libro mayor abierto y distribuido que pueda almacenar transacciones entre partes de manera verificable y permanente.

Una aplicación destacada es el intercambio de activos digitales, conocidos como criptomonedas. Las criptomonedas más conocidas y transadas son Bitcoin, Ethereum y Litecoin. Ofrecen, más allá de la transferencia de activos digitales, la ejecución de contratos inteligentes.

Los contratos inteligentes son esencialmente dos conceptos combinados. Una es la noción de software. Código frío y austero que hace lo que está escrito y se ejecuta para que el mundo lo vea. La otra es la idea de un acuerdo entre las partes. Son programas informáticos que facilitan, verifican y hacen cumplir la negociación y ejecución de contratos legales. Se ejecutan a través de transacciones de blockchain, interactúan con las criptomonedas y tienen interfaces para manejar la información de los participantes del contrato. Cuando se ejecuta en la cadena de bloques, un contrato inteligente se convierte en una entidad autónoma que ejecuta automáticamente acciones específicas cuando se cumplen ciertas condiciones.

Ethereum es la principal plataforma de computación distribuida pública basada en blockchain que proporciona un entorno de ejecución de contratos inteligentes dentro del contexto de una máquina virtual descentralizada, conocida como Ethereum Virtual Machine (EVM) [2, 3].

La máquina virtual EVM maneja el cómputo y el estado de los contratos y se basa en un lenguaje basado en una estructura de pila con un conjunto predefinido de instrucciones (códigos de operación) [3]. En esencia, un contrato es simplemente una serie de declaraciones de código de operación, que son ejecutadas secuencialmente por la máquina virtual EVM. La máquina EVM puede considerarse como una computadora global descentralizada en la que se ejecutan todos los contratos inteligentes. Aunque se comporta como una computadora gigante, es más bien una red de máquinas discretas más pequeñas en comunicación constante.

Los contratos inteligentes en Ethereum generalmente se escriben en lenguajes de alto nivel y luego se compilan en bytecode EVM. Los lenguajes de alto nivel son LLL (lenguaje tipo Lisp de bajo nivel), Serpent (un lenguaje de estilo Python), Vyper (un lenguaje de tipo Python), Bamboo (lenguaje basado en estados) y Solidity (un lenguaje estilo Javascript) [4]. LLL y Serpent se desarrollaron en las primeras etapas de la plataforma, mientras que Vyper y Bamboo se encuentran actualmente en desarrollo por la fundación Ethereum. El lenguaje más prominente y ampliamente adoptado es Solidity.

Solidity es un lenguaje de programación de alto nivel orientado a contratos. Su sintaxis es similar a JavaScript, está tipado de manera estática y admite herencia y polimorfismo, así como bibliotecas y tipos complejos definidos por el usuario.

En Solidity, los contratos se estructuran de manera similar a las clases en lenguajes de programación orientados a objetos. El código de contrato consta de variables y funciones que las leen y modifican, como en la programación imperativa tradicional.

El proceso de escritura de contratos seguros y de buen desempeño en Ethereum es una tarea difícil para los desarrolladores. Implica la aplicación de paradigmas de programación no convencionales, debido a las características inherentes de la ejecución de programas basados en blockchain. Además, los errores en los contratos desplegados pueden tener graves consecuencias, debido al acoplamiento inmediato del código del contrato y las transacciones financieras. Por lo tanto, es beneficioso tener una base sólida de diseño, patrones de código establecidos y probados que faciliten el proceso de escritura de código funcional y libre de errores, y herramientas que permitan analizar y detectar falencias de manera automática.

Un análisis de los contratos inteligentes existentes realizado por Bartoletti y Pompianu [5] muestra que las plataformas Bitcoin y Ethereum se centran principalmente en los contratos financieros. En otras palabras, la mayor parte del código del programa define cómo se mueven los activos (dinero). Por lo tanto, es crucial que la ejecución del contrato se realice correctamente. El manejo directo de los activos significa que las fallas tienen más probabilidades de ser relevantes para la seguridad y tienen mayores consecuencias financieras que los errores en las aplicaciones típicas.

Los incidentes, como el desbordamiento de valor en Bitcoin, o el ataque al proyecto DAO en Ethereum, causaron que una bifurcación dura de la blockchain anulara las transacciones maliciosas. Estos incidentes muestran que los problemas de seguridad se han utilizado con propósitos fraudulentos. Muchas de estas vulnerabilidades se pueden clasificar en tres grupos: lenguaje de programación, máquina virtual (EVM) y las peculiaridades propias de la blockchain, y pueden abordarse siguiendo las mejores prácticas para escribir contratos inteligentes seguros, que se encuentran dispersos en toda la comunidad Ethereum. Estas prácticas contienen principalmente información sobre los errores típicos que deben evitarse y la descripción de los enfoques de diseño y problemas.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En la búsqueda de asegurar los contratos inteligentes en la plataforma Ethereum se han adoptado diferentes enfoques, cubriendo aspectos como la semántica formal, patrones de seguridad y herramientas de verificación. De acuerdo con el tipo de enfoque abordado, se distingue entre enfoques de verificación y de diseño.

El objetivo de los enfoques de verificación es comprobar que los contratos inteligentes escritos en lenguajes de alto nivel existentes (como Solidity por mencionar alguno) cumplan con una política o especificación de seguridad, entre los que se pueden mencionar:

- Herramientas de análisis estático para la búsqueda automática de bugs [6, 7, 8, 9]
- Herramientas de análisis estático para la verificación automática de propiedades genéricas [10, 11, 12, 13]
- Frameworks para pruebas semi-automatizadas de propiedades específicas del contrato [14, 15, 16, 17]
- Monitoreo dinámico de propiedades de seguridad predefinidas [18, 19]

En contraste, los enfoques de diseño apuntan a facilitar la creación de contratos inteligentes seguros al proporcionar frameworks para su desarrollo: abarcan nuevos lenguajes que son más susceptibles de verificación, proporcionan una semántica clara y sencilla que es entendible por los desarrolladores de contratos

inteligentes o que permiten una codificación directa de las políticas de seguridad deseadas. Además, se incluyen trabajos que tienen como objetivo proporcionar patrones de diseño para contratos inteligentes seguros.

- Lenguajes de alto nivel [20, 21, 22, 23]
- Lenguajes intermedios [24]
- Patrones de seguridad [25]
- Herramientas [26]
- Ontologías [27, 28]

La línea de investigación presentada en este artículo busca plantear una estrategia para la detección de vulnerabilidades basada en ontologías, nutriéndose de ambos campos de estudio. Por un lado, utiliza un enfoque de verificación con las siguientes características:

- Lenguaje destino: lenguajes de alto nivel de la plataforma Ethereum (se selecciona Solidity como caso de estudio)
- Método de análisis: estático
- Garantías: búsqueda de bugs
- Grado de automatización: verificación automatizada

Por el lado de los enfoques de diseño, se toman como base los patrones de seguridad para lenguajes existentes [25] y la noción de ontologías para representar el dominio del problema, así como también, la idea de construir una representación intermedia para analizar los aspectos de seguridad.

3. RESULTADOS OBTENIDOS/ESPERADOS

El trabajo de investigación apunta a responder qué especificaciones tienen vulnerabilidades y presentar soluciones modeladas con las construcciones obtenidas como resultado de esta investigación y que serán utilizadas para mitigar los escenarios típicos de ataque. A continuación, se describe de manera sucinta algunos análisis comprendidos en la línea de investigación.

- Construcción de una representación intermedia que permita vincular el dominio del problema, representada por una ontología que comprenda los conceptos Ethereum, con el código fuente del contrato inteligente.
- Análisis estático de código fuente y matcheo de patrones de seguridad.
- Implementación de una herramienta que permita automatizar el proceso de análisis y detección de fallas de seguridad.

Los investigadores de esta línea continuarán con estudios en este campo con el objetivo de perfeccionarse en el área y realizar un seguimiento de nuevas amenazas, vulnerabilidades y ciberataques en materia de despliegue y ejecución de contratos inteligentes. También se planea seguir con los siguientes trabajos futuros:

- Generalización para otras plataformas de blockchain que soporten contratos inteligentes, tales como NEM, NEO, Cardano o Hyperledger.
- Obtención de código fuente a partir de bytecode EVM para aplicar el análisis que se propone en esta línea de investigación.
- Análisis de los aspectos de seguridad que surjan de la interoperabilidad con otras plataformas blockchain, también conocida como cross-chain smart contracts.

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 Doctor en Ingeniería Informática en la UNSL.

5. BIBLIOGRAFÍA

[1] Nakamoto, S., (2008), Bitcoin: Un Sistema de Efectivo Electrónico Usuario-a-Usuario, http://bitcoin.org/bitcoin.pdf

[2] Ethereum White Paper: https://github.com/ethereum/wiki/wiki/White-Paper

[3] Ethereum Yellow Paper: https://ethereum.github.io/yellowpaper/paper.pdf

[4] Solidity — solidity 0.4.18 documentation. [Online]. Available: https://media.readthedocs.org/pdf/solidity/develop/solidity.pdf

[5] Bartoletti, M. & Pompianu, L. (2017) *An empirical analysis of smart contracts: platforms, applications, and design patterns*. arXiv preprint arXiv:1703.06322.

[6] Luu, L., Chu, D.H., Olickel, H., Saxena, P., Hobor, A.: Making smart contracts smarter. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ACM (2016) 254–269

[7] Grishchenko, I., Maffei, M., Schneidewind, C.: A semantic framework for the security analysis of ethereum smart contracts. In: Proceedings of the 7th International Conference on Principles of Security and Trust (POST), Springer (2018)

[8] Zhou, E., Hua, S., Pi, B., Sun, J., Nomura, Y., Yamashita, K., Kurihara, H.: Security assurance for smart contract. In: New Technologies, Mobility and Security (NTMS), 2018 9th IFIP International Conference on, IEEE (2018) 1–5

[9] Nikolic, I., Kolluri, A., Sergey, I., Saxena, P., Hobor, A.: Finding the greedy, prodigal, and suicidal contracts at scale. arXiv preprint arXiv:1802.06038 (2018)

[10] Kalra, S., Goel, S., Dhawan, M., Sharma, S.: Zeus: Analyzing safety of smart contracts, NDSS (2018)

[11] Buenzli, F., Dan, A., Drachsler-Cohen, D., Gervais, A., Tsankov, P., Vechev, M.: Securify (2017) Available at http://securify.ch.

[12] Mythril Available at https://github.com/ConsenSys/mythril.

[13] Manticore Available at https://github.com/trailofbits/manticore.

[14] Hirai, Y.: Defining the ethereum virtual machine for interactive theorem provers. In: International Conference on Financial Cryptography and Data Security, Springer (2017) 520–535

[15] Amani, S., B´egel, M., Bortin, M., Staples, M.: Towards verifying ethereum smart contract bytecode in isabelle/hol. CPP. ACM. To appear (2018)

[16] Hildenbrandt, E., Saxena, M., Zhu, X., Rodrigues, N., Daian, P., Guth, D., Rosu, G.: Kevm: A complete semantics of the ethereum virtual machine. Technical report (2017)

[17] Bhargavan, K., Delignat-Lavaud, A., Fournet, C., Gollamudi, A., Gonthier, G., Kobeissi, N., Kulatova, N., Rastogi, A., Sibut-Pinote, T., Swamy, N., et al.: Formal verification of smart contracts: Short paper. In: Proceedings of the 2016 ACM Workshop on Programming Languages and Analysis for Security, ACM (2016) 91–96

[18] Grossman, S., Abraham, I., Golan-Gueta, G., Michalevsky, Y., Rinetzky, N., Sagiv, M., Zohar, Y.: Online detection of effectively callback free objects with applications to smart contracts. Proceedings of the ACM on Programming Languages 2(POPL) (2017) 48

[19] Cook, T., Latham, A., Lee, J.H.: Dappguard: Active monitoring and defense for solidity smart contracts

[20] OConnor, R.: Simplicity: A new language for blockchains. arXiv preprint arXiv:1711.03028 (2017)

[21] Pettersson, J., Edström, R.: Safer smart contracts through type-driven development

[22] Coblenz, M.: Obsidian: A safer blockchain programming language. In: Software Engineering Companion (ICSE-C), 2017 IEEE/ACM 39th International Conference on, IEEE (2017) 97–99

[23] Schrans, F., Eisenbach, S., Drossopoulou, S.: Writing safe smart contracts in flint

[24] Sergey, I., Kumar, A., Hobor, A.: Scilla: a smart contract intermediate-level language. arXiv preprint arXiv:1801.00687 (2018)

[25] W¨ohrer, M., Zdun, U.: Smart contracts: Security patterns in the ethereum ecosystem and solidity. (2018)

[26] Mavridou, A., Laszka, A.: Designing secure ethereum smart contracts: A finite state machine based approach. arXiv preprint arXiv:1711.09327 (2017)

[27] Ugarte, H.: BLONDiE https://github.com/hedugaro/Blondie

[28] Sitio oficial de la ontología EthOn: https://github.com/ConsenSys/EthOn

Esteganografía por sustitución múltiple, Implementación en Matlab.

Mg. Ing. Guillermo Sergio Navas[1]
Mg. Ing. Carlos Gustavo Rodríguez Medina[2]

Gabinete de Computación / Fac. de Ingeniería / Univ. Nacional de San Juan [1,2]

Av. Libertador Gral. San Martin 1109 (oeste) – San Juan
0264-4211700 (Int. 285[2] / 435[1])
snavas@unsj.edu.ar [1], grodriguez@unsj.edu.ar [2]

RESUMEN

Las técnicas para implementar esteganografía digital en portadores multimedia, para comunicación encubierta, deben responder a tres parámetros: Perceptibilidad, Capacidad y Detectabilidad. La tercera condición es imprescindible solo para casos especiales donde interesa que el portador con mensaje sea resistente al estegoanálisis. La premisa aquí es obtener las máximas capacidades, pero resulta en una solución de compromiso, ya que un aumento de Capacidad implica un incremento de perceptibilidad (y de detectabilidad).

Diversas técnicas se han propuesto, algunas son muy utilizadas hoy día, pero proveen *Capacidad* portadora más bien baja [1][2][3].

En el presente trabajo se expone una técnica, que llamamos "Sustitución Múltiple". En ella, para inyectar el mensaje, básicamente, se utilizan franjas de bits de los bytes correspondientes a cada canal de color de una imagen RGB. Actúa sobre las capas componentes inferiores consecutivas y completas de la imagen. No se aconseja utilizar más allá de las 4 primeras con este método.

Se logra así alcanzar un 37,5%[1] de capacidad en forma directa y segura, y hasta el 50% con elección de portador por análisis de ruido [4], lo cual es muy superior a otras técnicas en uso. Se propone también una variante resistente en cierto grado al estegoanálisis. Incrementos indirectos en la capacidad se obtienen por previa aplicación de algoritmos de compresión sobre el mensaje. Se suma importante aporte a la seguridad utilizando previos algoritmos de criptografía.

Se presenta también la implementación en Matlab, como programa piloto, que incluye la técnica propuesta y algoritmos de detección de ruido.

Palabras clave: Esteganografía, Sustitución Múltiple, Matlab.

CONTEXTO

El presente tratado expone una metodología esteganográfica, propuesta e implementada a nivel software, que incluye nueva técnica esteganográfica por Sustitución Múltiple y detección de ruido en imágenes; en el marco de un actual proyecto que fuera originado luego las tesis de posgrado *"Exploración de efectos esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos" y "Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganográficas"[5][6],* de Maestrías en Informática de la UNLaM.

Al proyecto lo conforma un equipo de investigación en la temática, en la Facultad de Ingeniería de la UNSJ.

1. INTRODUCCIÓN

La esteganografía es la disciplina en la que se estudian y aplican técnicas que permiten el ocultamiento de mensajes u objetos, dentro de otros, llamados portadores, de modo que no se perciba su existencia. Es una mezcla de artes y técnicas que se combinan para conformar la práctica de ocultar y enviar información sensible en un portador, de modo que pase inadvertido el hecho. La esteganografía juega un

[1] Estos extremos son siempre sin cobertura de *Detectabilidad*. Provee niveles muy altos de *capacidad* con *perceptibilidad* baja.

papel importante en la Seguridad de la información. Constituye una línea de investigación muy vigente, que se enmarca en lo concerniente a Seguridad Informática. Sus aplicaciones más comunes son: Comunicación encubierta y Protección de propiedad intelectual. Aquí nos enfocamos en la primera.

Los portadores preferidos son los archivos multimedia (imagen, video, sonido). El presente tratado se enfoca particularmente en portadores imagen BMP RGB[2] de 24 bits por la facilidad de analizar y exponer la técnicas.

Un imagen puede ser utilizada para ocultar, a la vista de intrusos cualquier mensaje u objeto software (archivo), codificándolo como sutiles cambios en los colores de los píxeles, en áreas poco significativas, que no deben ser percibidos por el ojo humano; de tal forma que el portador que contiene el mensaje, llamado *"Estegoportador"*, pueda ser transmitido, sin detectar el hecho. El receptor aplicará el proceso inverso (decodificación) para recuperar el mensaje oculto [5].

Las técnicas digitales empleadas para implementar esteganografía deben responder a tres parámetros: Perceptibilidad, Capacidad y Detectabilidad. La tercera condición resulta una exigencia en caso que se requiera resistencia al estegoanálisis[3]; son los casos, por ejemplo de aplicaciones militares, gubernamentales, espionaje, etc. Fuera de estos casos específicos, interesa más obtener un rendimiento o *Capacidad* lo más alta posible.

Ello resulta en una solución de compromiso, ya que un aumento de *Capacidad* incrementa la *perceptibilidad* (y *detectabilidad*).

En la presente propuesta, "*Sustitución Múltiple*", lo que prima es la obtención de elevada *Capacidad* en el portador, manteniendo un nivel aceptable de *Perceptibilidad*. Aunque una variante, que se mostrará, permite la cobertura a ataques comunes de estegoanálisis, es decir, cubre también *Detectabilidad*, lo cual se logra en detrimento de la *Capacidad*.

Diversas técnicas han sido propuestas, y son muy utilizadas actualmente, aunque proveen *Capacidad* relativamente baja, lo cual puede obligar a desmembrar el mensaje en varios portadores, con las consecuentes complicaciones. Así por ejemplo, en el dominio del espacio, se presentó en 2016 una mejora de BPHM (Block Pixel Hiding Method), llamada IPHM (Improved Pixel Hiding Method) [1], que logra duplicar la capacidad de la primera, la eleva a 25%, *pero posee una severa limitante*: un 5% de pérdida del mensaje, ello la limita a que los "archivos mensajes" solo puedan ser imágenes en tonos de grises, ya que ellas sí pueden ser interpretadas a pesar de la perdida. Se tiene también la conocida técnica LSB (Least Significant Bit) [2], quizá la más popular y una de las más utilizadas, que permite una máximo de 12,5%, ésta, además de su acotada *Capacidad,* tiene la desventaja de que es sumamente detectable. En el dominio de la frecuencia está la muy utilizada técnica basada en DCT (Discrete Cosine Transform) [3], aplicable a portadores JPG, la que suma propiedad de Robustez[4], pero con menor capacidad que a las otras.

Nota: la característica de Robustez se requiere solo para aplicaciones especiales, tales como watermarking y huellas digitales para copyright. Pero no es necesaria para comunicación encubierta.

En la técnica expuesta aquí se propone sustituir, de determinada manera, más de un bit del mensaje en cada byte del portador, múltiples de hecho, en general N, donde N=1...8. /

Se puede entender a una imagen RGB 24bits como compuesta de 8 capas, capas simples serían por cada canal de color, y compuestas las constituidas por los tres canales simultáneos. Una capa es una de sus imágenes binarias componentes. Cuando se aplica filtrado por capas lo que se obtiene es la vista de una ellas, o más de una solapadas (Ej. la capa 1, o la 2 y 3 juntas). Véase Fig. 1.

[2] Imágenes bitmap con tres canales de color, RGB proviene de sus siglas en inglés, Red, Green y Blue.

[3] El estegoanálisis es la disciplina dedicada al estudio de la detección de mensajes ocultos usando esteganografía.

[4] Propiedad esteganográfica que permite que el estegoportador sea sometido a compresiones, rotaciones, encuadres, etc. sin que se pierda el mensaje oculto.

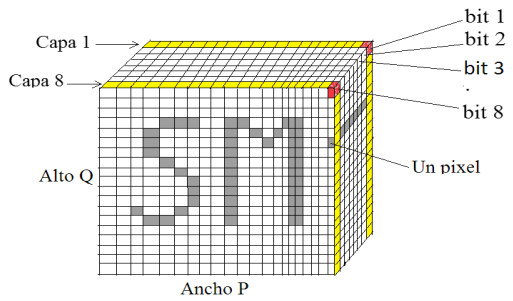


Fig. 1 - Modelo de representación de una imagen de 8 bits (un canal de RGB) de PxQ pixels.

El método de *Sustitución Múltiple* inserta el mensaje usando una o más capas de la imagen (una franja), utiliza N capas en general.

Usando una capa cualquiera de un solo canal de color (un mismo bit de todos los bytes de un color) estaríamos usando 1 bit por cada 3 bytes de la imagen (el pixel tiene 3 bytes: R G y B), ello daría una capacidad de 1/8x1/3= 0.04166, el 4.67%. Si lo hacemos en forma simultánea en los tres canales (capa compuesta), sustituiríamos 3 bits, 1 bit en cada una de la capas de color, con lo cual la capacidad se triplica, entonces Cap=0.04166x3=1/8=0.125, es decir se obtiene una utilización del 12.5%. La técnica de Sustitución Múltiple trabaja así, en los tres canales simultáneamente, es decir en las capas compuestas, de modo que por cada una (unidad de N) obtenemos 12.5% de capacidad. Si N=2 la capacidad es 25%, si N=3 es 37.5%. *El valor de N no solo hace referencia a la cantidad de capas a usar, sino también a su posición, contándose la primera como la capa 1* (Véase Fig. 1). Obviamente a medida que usamos más capas la *Perceptibilidad se incrementa*. Si N=1 se está manipulando la primer capa del portador.

Nótese que el máximo posible es N=8, máxima cantidad capas. Claro está, si bien es posible una sustitución con N=8, se estaría causando alteraciones en todas las zonas de la imagen, así la característica de *imperceptibilidad* no sería conseguida. En ese sentido, en el trabajo [5] se hizo un examen exhaustivo de la profundidad de bits, la posición de la capa usada y los efectos causados en las imágenes y se declaran los valores máximos permitidos para lograr una imperceptibilidad razonable, y según el nivel de ruido de la imagen ("categoría de imagen") [4].

Entonces, surge un interrogante lógico: ¿Cuántas capas y en qué posición se pueden utilizar sin notar alteraciones en el portador?

No hay una respuesta directa, y surgen propuestas que optimizan el proceso esteganográfico, analizando tipo de imagen, efectos sobre ella y características del portador [5].

Para tratar de dar una respuesta acotada y centrarse en la descripción de la técnica de interés, se puede decir que en primera instancia *un valor seguro para baja perceptibilidad es N=3*, más allá de ése dependerá del ruido del portador. N=4 se puede usar, pero para asegurar la imperceptibilidad se debe utilizar la técnica acompañada de algoritmos de selección de portador por nivel de ruido [4]. De aquí derivan las máximas capacidades asociadas a la técnica "Sustitución Múltiple": N=3, Cap=12.5% x 3 = 37.5%; y N=4, Cap = 12.5% x 4 = 50% (cada capa provee 12.5%).

Descripción y algoritmia de la técnica Sustitución Múltiple:

Una imagen color bitmap de 24 bpp está formada por tres canales de color, Rojo, Verde y Azul, cada uno de un byte de profundidad. Estas tríadas se disponen de forma consecutiva en la imagen. La conjunción de los tres canales forman cada pixel de la imagen.

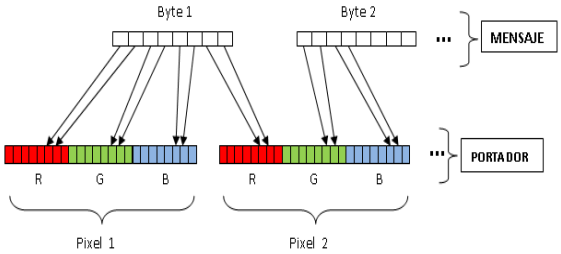


Figura 2. Implementación con N=2 partiendo desde la segunda capa. Cap. 25%. Cubre Detectabilidad.

Como se mencionó, en la técnica de Sustitución Múltiple, se reemplaza N bits de cada canal R, G y B; por sendos bits consecutivos del mensaje, en los pixeles necesarios de la imagen, hasta lograr ocultar todo el mensaje.

En la Fig. 2 se observa un diagrama para N = 2, donde se aprecia que cada dos bits correspondientes a los bytes del mensaje sustituye a 2 bits del canal R, análogamente para

los canales G y B. Luego se continúa ocultando los restantes bits del mensaje en cada uno de los siguientes pixeles del portador, de la misma forma, hasta que completar el mensaje.

Nótese que este caso ejemplo se están utilizando las capas 2 y 3, véase Fig. 1.

Observar que en este caso son necesarios *4* bytes del portador para ocultar *1* byte del mensaje, porque se están utilizando 2 capas; *capacidad recepción* resulta 25% (1/4x100).

Existe una relación de tamaños portador/mensaje a cumplir necesariamente, ella depende mucho de la técnica utilizada.

Respecto de otras técnicas en uso, la capacidad alcanzada aquí es muy superior, permite ocultar mayor cantidad de datos en un portador, incrementando la *capacidad de ocultamiento* en función del aumento del valor N.

La mayoría autores y desarrolladores considera que la utilización de más allá de la primer capa es inapropiada porque causa efectos visuales perceptibles. Hemos realizado rigurosos estudios, con ayuda herramientas software diseñadas al efecto, que demuestran que con *la utilización de hasta la tercera capa no se percibe a simple vista los efectos*, y esto es contando con la imagen original a los fines comparativos; difícilmente se detectarían alteraciones con N=3 sin usar herramientas, y de hecho, un intruso no dispondrá normalmente de la imagen original, por tanto le resultará imposible notar alteraciones. *La única condición para que resulte muy efectivo, el caso de N=3, es tratar de no utilizar portadores que tengan zonas de colores planos.*

En el ejemplo de la Fig. 2, se utilizaron dos capas, la segunda y tercera, *evitando la primera*. Se hace notar que de haber utilizado las 3 primeras capas (tres bits por byte de canal) el *efecto sobre perceptibilidad sería el mismo*, pero la capacidad sería del 37.5%. Entonces ¿Cuál es la ventaja de no usar la capa 1? Ella radica en que *aquí se tuvo en cuenta el factor Detectabilidad.* Los software de estegoanálisis[5] normalmente analizan la primera capa de una imagen, buscando alteraciones estadísticas por canal (son tres imágenes monocromo a procesar). Por una parte es porque así trabaja la mayoría de las técnicas esteganográficas más utilizadas, por otra, porque analizar varias capas apareja de un alto costo computacional, si se considera que el estegoanálisis suele ser realizado en línea mientras circulan las imágenes, y el tráfico es demasiado alto.

Así se ha intentado mostrar no solo el método de Sustitución Múltiple, sino *una variante* de él, que permite cubrir la detectabilidad. Obviamente, para casos especiales (aplicaciones militares, espionaje, etc.) el requerimiento es mayor, una Indetectabilidad muy alta; en tales casos se utiliza otra metodología, que escapa al alcance de este tratado.

Finalmente, un incremento indirecto de la capacidad se puede obtener aplicando al mensaje previamente un algoritmo de compresión, tal como el LZW (Lempel Ziv Welch). Y, por si un intruso lograra extraer el mensaje de un estegoportador es conveniente aplicar previamente al mensaje algún algoritmo de encriptación. Ambas cosas, compresión y encriptación, podrían ser parte de un software esteganográfico de usuario final que se desarrolle; pero si no lo provee, se puede usar, por ejemplo, 7-Zip, de alto ratio por LZMA (Lempel-Ziv-Markov) y LZMA2 y que puede cifrar los datos bajo el estándar AES de 256 bits [7].

Implementación Matlab:

Matlab es un software matemático que ofrece un entorno de desarrollo integrado (IDE) con un lenguaje de programación propio. Matlab es ampliamente utilizado en desarrollos científicos y en el ámbito académico.

Se ha optado por trabajar con este software, por la potencialidad que presenta para el procesamiento de imágenes, opera sobre ellas tratándolas como matrices. Por cada archivo imagen que se abre desde Matlab, se generan tres matrices, una para cada canal de color (Ver Fig 1). Se utiliza en conjunto con el Toolbox de procesamiento de imágenes. [6]

Se ha programado, según lo descrito en el apartado anterior, Fig.2, una función que implementa el algoritmo que aplica la Sustitución Múltiple. El formato es el siguiente:

SM(N) – Sustitución Múltiple

[5] El estegoanálisis representa la corriente contraria a la esteganografía, cuya finalidad es detectar mensajes ocultos en portadores.

N es el número de capas o bits por byte a sustituir en cada canal de color del portador, desde la 1ra.

SMV(N) – Variante que omite capa 1.

N es el número de capas o bits por byte a sustituir en cada canal de color del portador, desde la 2da.

Desde la línea de comando de Matlab se ejecuta función "*SM(N)*".Pide el ingreso de N, cantidad de bits por byte o capas que se desea usar del portador en sus canales de color.

La ejecución de la función solicita seleccionar el archivo de la imagen portadora, luego se selecciona el archivo de mensaje. Posteriormente, la función aplica el algoritmo de *Sustitución Múltiple* y presenta gráficamente una ventana con la imagen portadora original y otra con la imagen estegoportadora.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

La línea de investigación corresponde a la temática de *Esteganografía*, la que se enmarca en el área de Seguridad Informática. Los autores del presente trabajo han elaborado dos tesis de Maestría en este tema. Mantienen un proyecto que avanza en estudios y desarrollos de esteganografía. También han realizado publicaciones en diferentes congresos, particularmente en varias WICC. Los Trabajos realizados desde el año 2006 al presente, aportan resultados innovadores en el área.

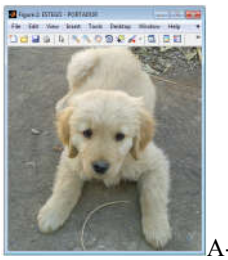
3. RESULTADOS OBTENIDOS

Se ha expuesto la técnica de Sustitución Múltiple y una variante que cubre detectabilidad; mencionando sus ventajas asociadas y la forma de aplicarla exitosamente. Exponiendo también el desarrollo de un software piloto implementado en Matlab. Con esa aplicación esteganográfica se puede realizar diversos ensayos con diferentes valores de "N". Ello permite realizar variadas pruebas y ver cómo influye en la imagen original la sustitución en distintas cantidades de capas. También se pueden aplicar realces de los efectos esteganográficos “no visibles”, con el fin de observar la porción y/o proporción ocupada por el mensaje. Adicionalmente se brindan algunos datos, como el porcentaje que ocupa el mensaje en el portador, y otros de interés.

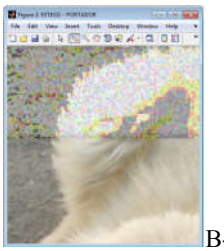
A- Portador bmp.

B- Mensaje bmp

Figura 3. Portador y mensaje utilizados.



A- Estegoportador.



B- Realce y Ampliación

Figura 4. Estegoportador con Realce de efectos y ampliación de imagen.

Para observar resultados con N=3, se selecciona un portador imagen BMP 24 bpp, de 400x418 pixeles y tamaño 490KB, Fig. 3-A.

Como mensaje se usó un archivo copia de pantalla de 42KB (portada del trabajo), Fig. 3-B. El mensaje podría haber sido de cualquier tipo (imagen, doc, ejecutable, etc.). El resultado, estegoportador, se muestra en la Fig. 4-A; observar que resulta ser visualmente idéntica a la original. A los fines demostrativos, se ha hecho un realce del efecto esteganográfico y ampliación de la zona, se aprecia el área ocupada por el mensaje oculto (que bajo condiciones normales no es visible), ello se expone en la figura 4-B.

4. FORMACIÓN DE RECURSOS HUMANOS

En la temática de Esteganografía, los integrantes del grupo de investigación vienen trabajando desde el año 2005. A fines de 2006, el primer autor de este trabajo defendió su Tesis de Maestría que trata del tema en profundidad. En octubre de 2015, el segundo autor realizó su defensa de Tesis de Maestría en Informática, que deriva y amplía la anterior, continuando el tratamiento del tema.

Actualmente se prosigue con el estudio, habiendo sumado un miembro más, que tiene en curso su tesis, en la misma maestría.

5. BIBLIOGRAFÍA

[1] Renza, Diego et al. *Método de ocultamiento de píxeles para esteganografía de imágenes en escala de gris sobre imágenes a color* [en línea]. Ingeniería y Ciencia, vol. 12, núm. 23, pp. 145-162. 2016. Universidad EAFIT. Colombia. Disponible en: http://www.scielo.org.co/pdf/ince/v12n23/v12n23a09.pdf

[2] Rodríguez, Gustavo; Navas, Sergio; *Esteganografía: Sustitución LSB 1 bit utilizando MatLab*. XVIII WICC, Universidad Nac. de Entre Ríos. 2016. Libro digital 1094p. Pág. 859-864, ISBN 978-950-698-377-2.

[3] Velasco Bautista et al. *Esteganografía en una imagen digital en el dominio DCT*. Científica [en línea] 2007. ISSN 1665-0654. Disponible en: https://www.redalyc.org/pdf/614/61411403.pdf

[4] Navas, G. Sergio; Rodríguez, Gustavo; Eterovic, Jorge; *Aplicación del filtro de Canny a la esteganografía digital*. Libro electrónico del XVI WICC. Ushuaia, Tierra del Fuego. 2014. Publicación digital ISBN 978-950-34-1084-4 .

[5] G. Sergio Navas. *Exploración de efectos Esteganográficos sobre portadores imagen de mapa de bits utilizando diferentes técnicas y algoritmos*. Argentina. Univ. Nacional de la Matanza – Escuela de Posgrado. Dic. de 2006.

[6] C. Gustavo Rodríguez M. *Estudio, análisis, desarrollo y propuestas de algoritmos para la selección óptima de métodos de sustitución en aplicaciones esteganogáficas*. Argentina. Univ. Nacional de la Matanza – Escuela de Posgrado. Nov. de 2015.

[7] Shubhi Mittal, Shivika Arora, Rachnma Jain, *PData security using RSA encryption combined with image steganography, Information Processing* (IICIP). 1st India International Conference on, pp. 1-5, 2016. Disponible en: https://ieeexplore.ieee.org/document/7975347

Estudio de Técnicas de Criptoanálisis

Castro Lechtaler, Antonio[1, 2]; Cipriano, Marcelo[1,3]; García, Edith[1], Liporace, Julio[1]; Maiorano, Ariel[1]; Malvacio, Eduardo[1]; Tapia, Néstor[1];

[1]Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática.
Facultad de Ingeniería del Ejército -FIE. Universidad de la Defensa Nacional - UNDEF
[2] CISTIC/FCE - Universidad de Buenos Aires.
[3] Departamento 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 proyecto de investigación propone estudiar las diferentes técnicas de Criptoanálisis, alcances y límites de las mismas, su marco de aplicación y su implementación mediante módulos interconectables mediante un framework.
Se orientarán las aplicaciones hacia los generadores de secuencias seudoaleatorias Stream Ciphers (*LFSR´s*[1], *NLFSR´s*[2], *CCG´s*[3] y *CA´s*[4]).
El desarrollo de herramientas de criptoanálisis permitirá la evaluación de algoritmos de cifrado para comprobar su robustez frente a los ataques que pudiera recibir, prestando sus funciones.
Entre las técnicas a abordar se encuentran las técnicas de Criptoanálisis Lineal [1], Diferencial [2-3], Algebraico, Guess-and-Determine [4] y Cube Attack [5], entre otras.
Este conjunto de herramientas posibilitará la realización de Análisis de Algoritmos de Cifrado, Generadores de Secuencias Seudoaleatorias, Primitivas Criptológicas, Protocolos de Seguridad, entre otros.

[1] Linear Feedback Shift Registers: registros de desplazamiento realimentados linealmente.
[2] Non Linear Feedback Shift Registers: registros de desplazamiento realimentados no linealmente.
[3] Clock Controlled Generators: generadores controlados por reloj.
[4] Cellular Automata: autómata celular.

Palabras Clave
Criptología, Técnicas de Criptoanálisis. Secuencias Seudoaleatorias.

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 Facultad de Ingeniería del Ejército (FIE) "Gral. Div. Manuel N. Savio", 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

En agosto de 2011, en una conferencia dada en el evento Techonomy[6] el CEO de Google Eric Schmidt sentenció que "la humanidad hasta el año 2003 había creado alrededor de 5 EB[5] de información" y luego "hoy en día, esa cantidad de información se genera cada 2 días y la tendencia va en aumento" [7-9].

[5] ExaByte: 2^{60} bytes de información.

La transmisión y almacenamiento de información requiere tomar una serie de medidas de protección manteniendo su confidencialidad, autenticidad e integridad. La falta total o parcial de medidas de seguridad se convierte en una amenaza latente. Cabe preguntarse entonces por la calidad y características de algoritmos criptográficos que se encargan de su protección y resguardo.

En la etapa de diseño de una primitiva criptográfica sus autores tienen en cuenta los ataques que éste pueda sufrir, a partir de las técnicas criptoanalíticas conocidas y a ellas lo someten, demostrando así su resistencia. Esta práctica ha llevado a los modernos algoritmos de cifrado a sortear amenazas, cuidadosa y eficientemente desarrolladas para atacarlos. Cada algoritmo, cada primitiva, cada protocolo debe ser atacado mediante una técnica adecuada a su estructura.

Por ello es que hay diferentes y variados ataques criptográficos y no existe uno que sirva para todos los algoritmos[6]. A su vez tampoco existe un único algoritmo que satisfaga todas las necesidades y requerimientos. Tal es así que suele llamarse a concurso para satisfacer la demanda de nuevos mecanismos que deberán mostrar su robustez y resistencia. Por ejemplo:

- el llamado en 1997 del *NIST* para escoger un nuevo algoritmo como estándar de cifrado llamado *AES* [10].
- El concurso europeo e-*Stream* en 2004, organizado por el *E-CRYPT* [11] 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, finalizado en 2012. Aunque aún no se da de baja al SHA-2 por no demostrar, hasta el momento, debilidades.
- La vigente competencia CAESAR7 (Competition for Authenticated Encryption: Security, Applicability, and Robustness) la cual ha emitido un portfolio de los algoritmos que han llegado a la final del certamen en 2018, entre los que se encuentran home v1, ordering addendum v1.3v1.4 v1.41 (Jérémy Jean, Ivica Nikolić, Thomas Peyrin, Yannick Seurin). [12].

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Para llevar adelante este proyecto de investigación se siguen las siguientes líneas de investigación y desarrollo:

- 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.
- Estudio, análisis y selección de los generadores de secuencias cifrantes.
- Relevamiento de los métodos criptoanalíticos que se analizarán.
- Estudio de técnicas criptográficas para la determinación del o los métodos de ataque adecuados a la estructura del algoritmo estudiado.
- Implementación de los métodos de criptoanálisis.

[6] El ataque por Fuerza Bruta puede aplicarse a cualquier algoritmo. El método consiste en probar todas las claves posibles, hasta encontrar la clave utilizada. Por las longitudes de las claves que se emplean actualmente, es posible que logre romper un criptosistema en varias veces la edad del universo. Es por ello que suele ser desaconsejable su implementación.

[7] CAESAR: Competition for Authenticated Encryption: Security, Applicability, and Robustness.

- Análisis de los resultados obtenidos.

3. RESULTADOS OBTENIDOS / ESPERADOS

Al realizar el estudio, análisis de las técnicas y herramientas de criptoanálisis, se persigue llevar adelante el diseño de aplicaciones criptográficas, evaluación y búsqueda de vulnerabilidades

El Criptoanálisis persigue:

- Obtener de la/s clave/s del cifrado.
- Encontrar patrones estadísticos en la salida del sistema.
- Desarrollar nuevas técnicas criptoanalíticas.
- Analizar el algoritmo de generación de la/s clave/s y estudiar su vulnerabilidad.

4. FORMACIÓN DE RECURSOS HUMANOS

Los docentes investigadores 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. Es por ello que 3 de ellos han demostrado su interés y se han sumado en calidad de colaboradores. En particular, el alumno Leiras, Facundo ha presentado su postulación en 2018 para la beca "Estímulo a las Vocaciones Científicas" (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones requeridas [13]. La misma le ha sido otorgada, iniciando en breve sus actividades respectivas.

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 y de la institución en la cual desarrollan sus actividades científico-docentes.

Por último y 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] 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.

[2] 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.

[3] 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.

[4] Pasalic, E.; "On Guess and Determine Cryptanalysis of LFSR-Based Stream Ciphers"; IEEE Transactions on Information Theory. Vol. 55 Ed.7º, 2009.

[5] 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.

[6] https://techonomy.com/ consultada el 27/2/19.

[7] https://readwrite.com/2010/08/04/goole_ceo_schmidt_people_arent_ready for_the_tech/. Consultada el 27/2/19.

[8] https://techcrunch.com/2010/08/04/schmidt-data/. Consultada el 27/2/19.

[9] Hilbert,m; López, p, “The World’s Technological Capacity to Store, Communicate, and Compute Information”. Science. 01 Apr 2011. Vol. 332, Issue 6025, pp. 60-65.
DOI: 10.1126/science.1200970.
[10] Daemen, J.; Rijmen, V.; “The Design of Rijndael: AES - The Advanced Encryption Standard”. Springer. New York. 2002.
[11] http://www.ecrypt.eu.org/stream/ Consultada el 10-3-18.
[12] https://competitions.cr.yp.to/caesar.html. Consultada el 10-3-18.
[13]http://evc.cin.edu.ar/informacion consultada el 23/2/2018.

Herramientas para la Evaluación de Algoritmos Criptográficos.

Liporace, Julio[1]; Cipriano, Marcelo[1,3] ; Malvacio, Eduardo; Estevez[2], Carlos; Fernández, Darío[1], García, Edith[1], López, Gabriel; Maiorano, Ariel[1] Ortega, Hugo[4]; Vera Batista, Fernando[1]

[1]Laboratorio de Investigación en Técnicas Criptográficas y Seguridad Teleinformática. Facultad de Ingeniería del Ejército (FIE), Universidad de la Defensa Nacional - UNDEF
[2] Instituto de Investigaciones Científicos y Técnicas para la Defensa – CITEDEF.-)
[3] Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes UNQ.
[4] Departamento Ciencias de la Computación. Facultad de Ciencias Exactas y Tecnología. Universidad Nacional de Tucumán

{marcelocipriano; dfernandez; verabatista}@est.iue.edu.ar
{edithxgarcia; jcliporace; maiorano; edumalvacio}@gmail.com
{hortega}@herrera.unt.edu.ar

RESUMEN

El aumento en el empleo de Internet en general y en particular de las tecnologías de Voz sobre IP, Teleconferencias, VideoStreaming, sistemas móviles y demás, han mostrado la necesidad de protegerlos mediante mecanismos criptográficos [1-2].

A los ya conocidos estudios de las propiedades matemáticas de los Generadores de Secuencias Pseudorandom generadas por Stream Ciphers [3] (algoritmos que involucran Linear Feedback Shift Registers [4], Non Linear Feedback Shift Registers [5]), Clock Controlled Generators y Autómatas Celulares, las Substitution Boxes creadas para los Block Ciphers [6] y los Generadores de Números Pseudoaleatorios que permiten evaluar propiedades criptológicas [7-8], se le debe agregar la posibilidad de la existencia de "puertas traseras" o "back-doors" presentados en recientes trabajos han sorprendido a la comunidad criptológica [9-10].

Es por ello que se persigue la creación de herramientas que permitan automatizar tales análisis y la realización de pruebas, para llevar adelante el estudio de manera veloz y eficiente. Al ser sometidos a este escrutinio en profundidad, los algoritmos podrán ser calificados de acuerdo a las propiedades que manifiesten. De esta forma, el usuario podrá decidir acerca de su uso, de acuerdo al nivel de seguridad que se precise y el que el algoritmo finalmente ofrezca.

Muchas propiedades criptológicas quedan ocultas detrás de las líneas de programación o en la complejidad matemática que compete a estos mecanismos [11]. Suele ocurrir que las explicaciones técnicas más profundas no se abordan en detalle o son parcialmente expuestas [12]. Los investigadores deben profundizar en cada algoritmo y mediante su estudio, deducir sus propiedades.

Palabras Clave
Criptología, Criptoanálisis. Stream Ciphers.

CONTEXTO

"Herramientas para la Evaluación de Algoritmos Criptográficos" (HEAC), es un proyecto de investigación y desarrollo que se presenta dando cumplimiento a la Resolución Rectoral 154/18 en el marco del Programa de Acreditación y Financiamiento de Proyectos UNDEFI, perteneciente a la Universidad de la Defensa Nacional (UNDEF). Luego de un

período de evaluación, fue seleccionado con el Nro. 340 y recibe financiamiento de la Universidad.
La investigación se lleva adelante en el contexto de la carrera de grado de Ingeniería en Informática y el posgrado en Criptografía y Seguridad Teleinformática. Ambos se dictan en la Facultad de Ingeniería del Ejército "Gral. Div. Manuel N. Savio" (FIE), perteneciente a la Universidad de la Defensa Nacional (UNDEF). Allí se llevan adelante tareas de I+D+i por parte del Grupo de Investigación en Criptología y Seguridad Informática (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

En los últimos años los métodos para el diseño de algoritmos seguros han tenido un gran avance e impulso a nivel mundial. Basta recordar, entre otros:

- Los llamados en 1997 del NIST para escoger un nuevo algoritmo como estándar de cifrado llamado AES.
- El inicio en el año 2000 del proyecto del gobierno japonés llamado CRYPTREC, creado para evaluar y recomendar técnicas criptográficas para uso gubernamental e industrial. La primera "Lista de Algoritmos Recomendados" fue publicada en 2003, con una revisión en el año 2013.
- El concurso europeo e-Stream organizado por el European Network of Excellence in Cryptology (E-CRYPT) que se llevó adelante entre el 2004 y 2008 con el propósito "Identificar nuevos algoritmos de cifrado en cadena". De él surgieron 7 algoritmos: 4 pertenecientes al Portfolio correspondiente a software y 3 pertenecientes al Portfolio correspondiente a hardware.
- SHA-3 (Secure Hash Algorithm) el nuevo estándar hash, a partir de conocer a través de NIST en 2012 el vencedor del certamen (KeccaK), aunque aún puede seguir usándose SHA-2.
- PHC (Password Hashing Competition) que corrió entre 2013 al 2015, recomendando finalmente al algoritmo ARGON2 para llevar adelante el hash de claves.
- Y la aún vigente competencia CAESAR (Competition for Authenticated Encryption: Security, Applicability, and Robustness) la cual ha emitido un portfolio de los algoritmos que han llegado a la final del certamen en 2018, entre los que se encuentran home v1, ordering addendum v1.3v1.4 v1.41 (Jérémy Jean, Ivica Nikolić, Thomas Peyrin, Yannick Seurin).

A su vez, el reconocimiento a través de normas de muchos algoritmos criptográficos fue en aumento, como ser las ISO (International Organization for Standardization). Esta es una organización internacional no gubernamental, formada por organismos de estandarización de 163 países miembros y el IEC (International Electrotechnical Commission), es una organización de normalización que actúa sobre temas eléctricos, electrónicos y tecnologías asociadas. Está formada por los organismos nacionales de normalización, al igual que la International Organization for Standardization, en las áreas mencionadas de los 83 países miembros. Muchas normas se desarrollan en común con

ISO, por ello se las llama 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.
- 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.
- 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.

Este renacimiento mundial por la búsqueda de nuevos algoritmos resulta bienvenido. Pero por sí sólo resulta insuficiente a la hora de establecer parámetros criptográficos seguros.

Cada algoritmo, cada primitiva, cada protocolo debe ser atacado mediante una técnica adecuada a su estructura. Es por ello que en la actualidad no existe una única modalidad general de criptoanálisis. Los autores ya enfocan el diseño de algoritmos de cifrado para que resulten indemnes ante los ataques conocidos. Así el impacto del Criptoanálisis resulta de peso puesto que los algoritmos criptológicos, los protocolos y también los tamaños de las claves entre otros, son seleccionados basándose en él.

Aunque el objetivo fundamental del criptoanálisis es hallar las vulnerabilidades en uno o varios aspectos de la seguridad de los algoritmos criptológicos, implícita o explícitamente, no se extingue allí su alcance. Sino que también resulta ser la herramienta fundamental a la hora de establecer la fortaleza de los mismos. No debería adoptarse ningún algoritmo si no se prueba que éste ha sido sometido a distintas técnicas de criptoanálisis y ha resistido eficientemente a las mismas

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

Para llevar adelante la investigación y el desarrollo del proyecto, se proponen las siguientes líneas de acción:

- Investigación bibliográfica acerca de las técnicas de ataque criptoanalíticas conocidas.
- Seguimiento y/o asistencia a los congresos y workshops de relevancia y referencia del tema abordado, tanto nacionales como internacionales, acerca del progreso del criptoanálisis de los algoritmos estudiados., como así también la presentación de nuevas herramientas para la búsqueda de vulnerabilidades.
- Determinación de la metodología que se implementará a lo largo del desarrollo del proyecto.
- Elaboración de módulos para las pruebas algorítmicas.
- Aplicación de las herramientas a primitivas criptográficas en particular.

3. RESULTADOS OBTENIDOS / ESPERADOS

Es por ello que el proyecto persigue la creación de herramientas informáticas que permitan automatizar algoritmos de prueba y análisis que posibiliten la detección de debilidades y/o vulnerabilidades criptoanalíticas que pudieran ser explotadas por un atacante, mediante técnicas de criptoanálisis. Y todo ello

llevado delante de forma veloz y eficiente.
Se someterá a un análisis profundo a los algoritmos en cuestión, para así poder ser calificados de acuerdo a las propiedades que manifiesten. De esta forma, el usuario podrá decidir acerca de su uso, de acuerdo al nivel de seguridad que se precise y la que el algoritmo finalmente ofrezca.
Muchas de las propiedades criptológicas quedan solapadas u ocultas detrás de las capas del lenguaje de programación o el diseño informático del mismo, como así también, en la maraña matemática que encierran estos mecanismos. Suele ocurrir que las explicaciones técnicas más profundas no son abordadas en detalle y son parcialmente expuestas. Los investigadores deben profundizar en cada algoritmo y mediante su estudio, deducir sus propiedades.
Asimismo al considerar además el ciclo de vida de cada algoritmo, se podrá observar el avance o progreso de los ataques a los que se lo ha sometido. Muchas veces esos ataques conllevan un éxito y se van observando vulnerabilidades. La mayoría de ellas son resueltas con el advenimiento de versiones mejoradas de los mismos. Esta permanente evolución debe ir acompañada de la actualización de pruebas y demás procesos que permitan evaluar su seguridad.
Los objetivos particulares que presenta el proyecto son:

- 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

Este proyecto se lleva se lleva adelante en conjunto con la Universidad Nacional de Tucumán, mediante la colaboración de uno de sus docentes investigadores, tal como figura en el parágrafo correspondiente a los autores.
Los docentes investigadores 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. Es por ello que 3 de ellos han demostrado su interés y se han sumado en calidad de colaboradores. En particular, el alumno Leiras, Facundo ha presentado su postulación en 2018 para la beca "Estímulo a las Vocaciones Científicas" (EVC) otorgadas por el Consejo Interuniversitario Nacional (CIN) por encuadrarse en las condiciones requeridas [13]. La misma le ha sido otorgada, iniciando en breve sus actividades respectivas.
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 y de la institución en la cual desarrollan sus actividades científico-docentes.
Por último y 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] Wu, H. "Related-Cipher Attacks" International Conference on Information and Communication Security. (ICICS 2002), Lecture Notes in Computer Science 2513, R. Deng, S. Qing, F. Bao and J. Zhou (Eds.), pp. 447-455, Springer-Verlag, (2002).

[2] Souradyuti, P.; Preneel, B. “On the (In)security of Stream Cyphers Based on Arrays and Modular Addition”, Advances in Cryptology - proceedings of ASIA-CRYPT 2006, Lecture Notes in Computer Science 4284, pp. 69-83, Springer-Verlag, (2006).
[3] Fischer, W.; Gammel, B.; Kniffler, O.; Velton, J. “Differential Power Analysis of Stream Ciphers,” Topics in Cryptology-CT-RSA 2007, Springer-Verlag, LNCS, Vol. 4377, pp. 257–270. (2007).
[4] Masoodi, F.; Alam, S; Bokhari, M. “An Analysis of Linear Feedback Shift Registers in Stream Ciphers”. International Journal of Computer Applications 46, pp. 46-49, May 2012. Published by Foundation of Computer Science, New York, USA. (2012).
[5] Biryukov, A.; Shamir, A. “Cryptanalytic time/memory/data tradeoffs for stream ciphers”. In T. Okamoto, editor, ASIACRYPT, volume 1976 of Lecture Notes in Computer Science, Pp 1–13. Springer. (2000).
[6] Biryukov, A. “Block ciphers and stream ciphers: The state of the art”.IACR Cryptology ePrintArchive, Pp: 1–22. (2004).
[7]Albrecht, M.; Cid, C. “Algebraic techniques in differential cryptanalysis”. InFastSoftware Encryption, Pp 193–208. Springer-Verlag, (2009).
[8] Carlet, C.; Feng, K. “An infinite class of balanced functions with optimal algebraicimmunity, good immunity to fast algebraic attacks and good nonlinearity”. In14th InternationalConference on the Theory and Application of Cryptology and Information Security, Advancesin Cryptology (ASIA CRYPT 2008), volume 5350 of Lecture Notes in Computer Science, Pp 425–440. Springer-Verlag, (2008).
[9] Calderini M., Sala M.:”Elementary abelian regular subgroups as hidden sums for cryptographic trapdoors”. arXiv:1702.00581 (2017).
[10] Brunetta C., Calderini M., Sala M.: “Algorithms and bounds for hidden sums in cryptographic trapdoors”. arXiv:1702. 08384 (2017).
[11] Civino, R., Blondeau, C. & Sala, M. “Designs, Codes and Cryptography”. Pp 225-247. https://doi.org/10.1007/ s10623-018-0516-z. (2019)
[12] Aiston J. “Ring Theoretic Key Exchange for Homomorphic Encryption”. In: Arai K., Bhatia R. (eds) Advances in Information and Communication. FICC 2019. Lecture Notes in Networks and Systems, vol 70. Springer, Cham. Springer Nature Switzerland AG. (2019).
[13]http://evc.cin.edu.ar/informacion, página consultada el 23/2/2019.

Integridad de archivos y su distribución en una red Blockchain

Sebastian Ponti, Lautaro Nahuel De León, Patricia Bazán
LINTI Facultad de Informática UNLP
sebaponti@gmail.com, laudleon@gmail.com, pbaz@ada.info.unlp.edu.ar

RESUMEN

En la actualidad, los métodos para garantizar la integridad de archivos tienen como problemática que el respaldo de dicha característica se apoya en un servidor central, con lo cual, es relativamente simple que alguien con malas intenciones pueda realizar alteraciones de estos datos, sin dejar rastros. En el año 2008, se publicó un paper titulado "Bitcoin: Un Sistema de Efectivo Electrónico Usuario-a-Usuario" [1] el cual explica el funcionamiento de un sistema monetario electrónico totalmente descentralizado, es decir, sin intervención de una entidad tercera que valide la autenticidad de las transacciones. El principio básico de este sistema es justamente distribuir la información entre todos los nodos dentro de la red de forma tal que cada uno de ellos mantenga en todo momento el mismo historial de transacciones realizadas y donde las mismas sean completamente inmutables, a fin de mantener la buscada integridad y confiabilidad de estos datos.
En este artículo se enuncia una alternativa de solución al problema mencionado a través de una aplicación concreta utilizando la plataforma Ethereum [2], la cual es una implementación de una blockchain que permite realizar cómputo distribuido y seguro sobre ella por medio de contratos inteligentes [3].

CONTEXTO

El presente trabajo se encuentra inserto en un área de innovación que parte originalmente del estudio de los sistemas distribuidos -particularmente los descentralizados-, pero que en los últimos años, dado el vertiginoso crecimiento de los fundamentos propios que caracterizan a este tipo de arquitecturas, las blockchain, ha divergido en una línea de investigación propia y sumamente fértil, con un vasto potencial para muchas aplicaciones.

1. INTRODUCCION

Desde el surgimiento de Internet en 1990 y el advenimiento de la descarga de archivos a través de esta red, siempre ha sido un problema garantizar la integridad de los mismos, es decir, poder probar mediante un método riguroso que estos no son alterados en ningún momento. Estas potenciales modificaciones pueden ser aleatorias, por ejemplo, debido a fallas del medio de transmisión, o bien, porque un agente malintencionado introduce un cambio indebido.
Dentro del marco de Internet, uno de los primeros algoritmos utilizados para verificar la integridad de los archivos transferidos fue MD5 [4], el cual pertenece a la familia de funciones de *hash* de una sola vía (*one-way function*). Este tipo de funciones tiene como objetivo asociar cada elemento del dominio que, en el presente caso, está caracterizado por el conjunto de archivos posibles (el cual es infinito) con una cadena pseudo-aleatoria y fija de caracteres, la cual recibe el nombre de huella digital. La particularidad de las funciones de una sola vía es que el costo computacional que se invierte para obtener dicha huella digital a partir de un archivo - o cualquier fuente en general- es muy barato, mientras que el proceso inverso es virtualmente imposible. Esta característica las hace especialmente idóneas para la verificación de integridad de archivos sobre redes de datos:

genéricamente, si en un nodo cualquiera de la red se computa la huella digital y se deja accesible de forma pública, entonces cualquier otro nodo de la misma red podría corroborar que el archivo transferido fue recibido de forma intacta computando la misma función sobre él, y luego, verificando que el resultado sea igual a la huella generada por el nodo origen; de lo contrario, se infiere que hubo algún tipo de alteración sobre él.

Aun en el caso de disponer de una función de una sola vía relativamente segura -en 2005 se demostró que MD5 era insegura [5]-, muchas veces el mayor problema recae en la arquitectura, particularmente, aquellas en la que la seguridad se concentra en un único punto o servidor. Además de los problemas de rendimiento conocidos, en donde un número reducido de servidores podría colapsar ante una gran cantidad de peticiones por parte de los usuarios o clientes [6], existen también problemas de seguridad asociados. Si, por ejemplo, un atacante consigue tomar control del/los servidor/es del sistema, entonces cualquier información alojada en ella dejaría de ser automáticamente confiable, en particular, las huellas digitales publicadas de los archivos ubicados allí. Al administrar el servidor, un agente malintencionado podría cambiar los archivos, recomputar sus huellas y volver a publicarlas al exterior y todo el circuito parecería confiable, pero sin embargo, el agente ha conseguido romper un eslabón de confiabilidad y eso podría ser muy costoso para el proceso o negocio involucrado.

Dadas las problemáticas de las redes centralizadas basadas en arquitecturas cliente-servidor, una variante desarrollada en pos de mitigar este inconveniente fue pensar las redes sin la noción típica de "servidor central", sino más bien, pensar en redes donde cada nodo participante de ella pueda requerir información o trabajo como así también suministrar información o trabajo. Cabe destacar que hay distintos niveles de descentralización; por ejemplo, lo que se describió previamente es una instancia de red descentralizada pura dado que todos los nodos actúan como clientes y servidores a la vez. Sin embargo, esto no necesariamente se da así en la realidad. En otros casos la red puede ser mixta, esto es, una red con un número equilibrado y bien distribuido -ya sea lógica o físicamente- de servidores y clientes. Estas características son cruciales para el presente trabajo puesto que la confianza ahora no dependerá solamente de un único punto sino de muchos otros, y esto, disminuirá exponencialmente la probabilidad de obtener todo el control de la red con el objeto de anunciar información espuria sobre ella.

Aproximadamente a partir de la década del 80 se han desarrollado una gran cantidad de ideas abogando por las bondades de la descentralización [7] [8] [9], llegando hasta el mencionado año 2008 en donde se publica el trabajo de Satoshi Nakamoto (pseudónimo) sobre la red Bitcoin, implementando el primer modelo completo de lo que hoy se conoce como Blockchain. En esta arquitectura, todas las transacciones son públicas y distribuidas, y una vez validadas, no pueden revertirse ni alterarse. Esto se consigue agrupando a las mismas en bloques, en donde cada una de estas últimas conforman los eslabones de una cadena en la cual cada uno se identifica con una huella digital que depende parcialmente de las huellas de los bloques previos, protegiendo así a la cadena completa de modificaciones, dado que, de lo contrario, la cadena de huellas quedaría completamente inválida [10] [11].

Por otro lado, con el objeto de evitar que cualquier nodo publique transacciones falsas, se exige que cada bloque lleve consigo una prueba de trabajo (*PoW*), en donde los nodos mineros o validadores de transacciones y bloques resuelvan un desafío computacional muy costoso de llevar a cabo pero muy sencillo de verificar con el propósito de que, justamente,

un nodo demuestre interés genuino en validar, y por ende, asegurar el buen funcionamiento de la red por medio de su inversión de trabajo [12].

Dado que Bitcoin únicamente sirve para transacciones monetarias, en el año 2014, sale a la luz la plataforma Ethereum, la cual implementa también una Blockchain con los mismos fundamentos que la de Bitcoin, pero con la gran diferencia y ventaja de que puede albergar fragmentos de código programables por los usuarios llamados "contratos inteligentes". Estos, al igual que una transacción, se pueden ejecutar y validar en cualquier nodo de la red, otorgando cómputo descentralizado.

En pos de resolver el problema previamente explicado en este artículo, se desarrollará una pequeña aplicación cuyo objetivo será el poder certificar archivos y distribuir dicha certificación -una huella digital del mismo, en términos simples- en la red Ethereum a través de los mencionados contratos inteligentes con el fin de mitigar considerablemente el riesgo de ataque y malversación del servicio que brinda la certificación.

La arquitectura de la aplicación contará con las siguientes componentes e interacciones:

- Frontend: una aplicación que brinde las operaciones necesarias al usuario como para listar, verificar y dar de alta certificaciones de archivos.
- Servidor de archivos: lugar donde residirán los archivos. Debería proveer por archivo una huella y una identificación unívoca y anónima.
- Blockchain (Backend): el corazón de la aplicación, en donde se persistirán y distribuirán las huellas de los archivos.

Lo que se intenta describir son los tres componentes básicos, el *frontend*, el servidor de archivos y el *backend* o blockchain, el cual necesariamente estará distribuido y replicado por su naturaleza intrínseca. Habrá interacciones bidireccionales entre el *frontend* y el servidor de archivos y entre el primero y la blockchain para las distintas operaciones a especificar. Como se ha dicho anteriormente, opcionalmente el servidor de archivos podría estar replicado y las aplicaciones clientes también, si en futuro, por ejemplo, fueran aplicaciones móviles.

A modo de ejemplo, se describe una de las operaciones del sistema, la de alta de una certificación, por medio de un diagrama de secuencia (ver figura 1).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

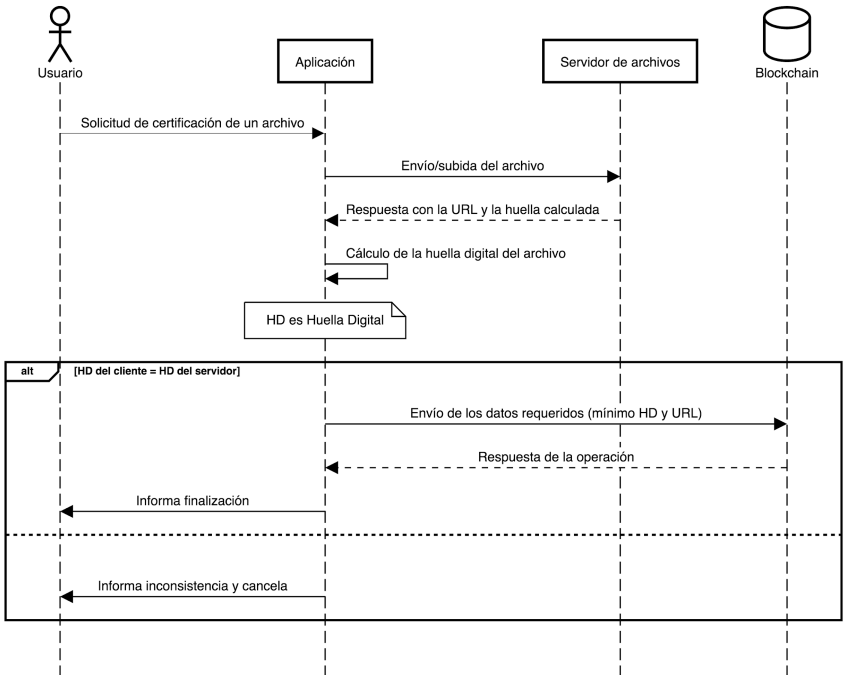


Figura 1: Diagrama de secuencia de alta de una certificación

El trabajo pone en manifiesto, por un lado, los conceptos de sistemas distribuidos, y particularmente, aquellos referidos a Blockchain y Ethereum: transacciones, consenso distribuido y contratos inteligentes. Por otro lado, se atacan cuestiones propias e inherentes al desarrollo de este tipo de aplicaciones, que interactúan con componentes desplegados en la Blockchain, como por ejemplo, diseño de la infraestructura, testing, análisis y evaluación de las diversas tecnologías presentes para llevar a cabo dichas tareas [13] [14] [15] [16] [17].

3. RESULTADOS OBTENIDOS/ESPERADOS

El resultado que intenta otorgar este trabajo es la capacidad de corroborar la existencia e integridad de un archivo en un momento determinado -prueba de existencia-, utilizando un medio de alojamiento en el cual no puede ser alterado o destruido.

Asimismo se hace foco en aportar un nivel de confianza y seguridad para el usuario que decida corroborar que, por ejemplo, evidencias fotográficas en un proceso de auditoría no fueron falsificadas luego de haberlas subido a un servidor centralizado.

4. FORMACION DE RECURSOS HUMANOS

La línea de investigación que se describe en el presente trabajo involucra el desarrollo de aplicaciones descentralizadas, la cual propone un nuevo paradigma a la arquitectura clásica que se conoce hoy en día en las aplicaciones orientadas a sistemas distribuidos. Debido a que la línea de investigación se estudia en una época de maduración temprana de la tecnología blockchain, el número de profesionales e investigadores involucrados en la misma es un grupo reducido.

En el presente trabajo se describe la línea de investigación que dio lugar a que en el año 2018 se formule la primera propuesta para tesina de grado haciendo uso de la tecnología blockchain como eje principal.

5. BIBLIOGRAFIA

[1] Satoshi Nakamoto. Bitcoin. A Peer-to-Peer Electronic Cash System, 2008.

[2] Vitalik Buterin et al. A next-generation smart contract and decentralized application platform. white paper, 2014.

[3] Nick Szabo. Formalizing and securing relationships on public networks. First Monday, 2(9), 1997.

[4] Ronald Rivest. The md5 message-digest algorithm. 1992.

[5] Xiaoyun Wang and Hongbo Yu. How to break md5 and other hash functions. In Annual international conference on the theory and applications of cryptographic techniques, pages 19–35. Springer, 2005.

[6] J. E. Neilson, C. M. Woodside, D. C. Petriu, and S. Majumdar. Software bottlenecking in client-server systems and rendezvous networks. IEEE Transactions on Software Engineering, 21(9):776–782, Sep 1995. ISSN 0098-5589. doi: 10.1109/32.464543.

[7] David L. Chaum. Untraceable electronic mail, return addresses, and digital pseudonyms. Commun. ACM, 24(2):84–90, February 1981. ISSN 0001-0782. doi: 10.1145/358549.358563. URL http://doi.acm.org/ 10.1145/358549.358563.

[8] David Chaum. Computer systems established, maintained and trusted by mutually suspicious groups /. Jan 1982.

[9] David Chaum. Blind signatures for untraceable payments. In David Chaum, Ronald L. Rivest, and Alan T. Sherman, editors, Advances in Cryptology, pages 199–203, Boston, MA, 1983. Springer US. ISBN 978- 1-4757-0602-4.

[10] Stuart Haber and W. Scott Stornetta. How to time-stamp a digital document. J. Cryptol., 3(2):99–111, January 1991. ISSN 0933-2790. doi: 10. 1007/BF00196791. URL http://dx.doi.org/10.1007/BF00196791.

[11] Dave Bayer, Stuart Haber, and W. Scott Stornetta. Improving the efficiency and reliability of digital time-stamping. pages 329–334, 1993.

[12] Adam Back. Hashcash - a denial of service counter-measure. August 2002.

[13] Ethereum. Solidity. https://solidity.readthedocs.io/, 2019.

[14] Node.js. Express.js. https://expressjs.com/, 2018.

[15] Facebook. React. https://reactjs.org/, 2018.

[16] Web3. Web3. https://web3js.readthedocs.io/en/1.0/, 2018.

[17] Metamask. Metamask. https://metamask.io/, 2018.

Investigación en ciberseguridad: Nuevos desafíos para adaptarse a nuevos paradigmas

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 de manera ininterrumpida una línea de investigación en seguridad informática, inicialmente, y en ciberseguridad en la actualidad[1]. Importantes proyectos como el establecimiento del CSIRT académico CERTUNLP[2] y de la autoridad de certificación PKIGRID UNLP[3], así como el afianzamiento del equipo de CTF SYPER[4] plasmado en varios premios obtenidos, son ejemplos concretos de implementaciones que plasman las actividades de investigación que se llevan a cabo.

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 sumando nuevas temáticas como la implementación de monitoreo de seguridad inteligente y el análisis de nuevas amenazas en distintas tecnologías.

Las actividades de investigación presentadas se retroalimentan con el habitual compromiso de trasladar los resultados 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: Ciberseguridad, seguridad inteligente, CTFs, Seguridad en aplicaciones, IoT

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"[5] 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.

1. INTRODUCCIÓN

En un mundo cada vez más conectado y con mayor penetración de la tecnología los problemas se multiplican y profundizan, ya que la superficie de ataque crece día a día y las tecnologías continúan mutando y evolucionando. Estos cambios tecnológicos generan que si bien conceptos tradicionales de seguridad de la información como firewalls, los sistemas de detección de intrusiones, los monitores de seguridad de red o los antivirus

continúen siendo necesarios, los mismos tienen que ser complementados con nuevas técnicas y herramientas que enriquezcan la gestión de la ciberseguridad.

Para citar un ejemplo, si uno analiza tráfico de red se notarán grandes diferencias con lo que ocurría hace unos años; en la actualidad se encontrarán grandes volúmenes de información y mucho uso de protocolos de cifrado. Esto lleva a la necesidad de cambiar de paradigma y se percibe que entre las principales tendencias de ciberseguridad se encuentran una percepción de escasez de recursos capacitados y un cambio en el enfoque hacia la detección y respuesta casi en tiempo real[6].

Estos desafíos requerirán nuevos tipos de habilidades en ciencia de datos y análisis. Dado que resulta imposible mitigar todas las amenazas para evitar su concreción, se debe cambiar el enfoque de la ciberseguridad. El esfuerzo debe dejar de concentrarse únicamente en la protección y prevención, para conducirlo más equitativamente a la detección y respuesta.

El término "seguridad inteligente" que se ha comenzado a difundir en los últimos años, hoy en día toma empuje. Los productos de seguridad inteligentes son herramientas que combinan aspectos del aprendizaje automático y de la inteligencia artificial con las aplicaciones de seguridad tradicionales. Estas cuentan con mayor capacidad para adaptarse a nuevas amenazas y asegurar nuevos tipos de aplicaciones. Tienen la capacidad de identificar comportamiento sospechoso, cambios en el malware existente y monitorear contínuamente el tráfico de red para descubrir eventos anómalos.

En los últimos años se han propuesto enfoques basados en tecnologías de aprendizaje automático para detectar malware, aplicando análisis estático y dinámico. Ambos enfoques presentan algunas desventajas. En el caso del análisis estático, el cifrado y las técnicas de ofuscación dificultan la extracción de características del malware. Por otro lado, el análisis dinámico requiere simular el entorno de operación del malware y observar su comportamiento durante un período largo de tiempo, lo cual resulta costoso.

La evolución del aprendizaje automático tiene a la ciberseguridad como nuevo hogar dentro de la transformación digital, ya que las herramientas lo utilizan junto a las prácticas de monitoreo contínuo para ayudar a defenderse, detectar y remediar las amenazas cibernéticas[7].

Más allá de todos los esfuerzos que se lleven a cabo en pos de la seguridad proactiva, los incidentes seguirán ocurriendo, con lo cual es importante gestionarlos en forma eficiente, de manera de minimizar el impacto de los mismos tanto para la organización, como para los usuarios y la red toda. A través del uso de herramientas que automaticen tareas y procedimientos y que se integren con múltiples fuentes de información que detectan problemas, y de los cuales puedan nutrirse posibilitará una gestión más eficiente de los incidentes de seguridad y permitirá mejorar los niveles de madurez de seguridad en las organizaciones.

En términos de mejorar los niveles de seguridad de las organizaciones, la seguridad de los sistemas críticos constituye una de las principales preocupaciones. Ello debe ser considerado en todo el ciclo de vida de los sistemas, desde su desarrollo a su puesta en producción. Es necesario tener en cuenta por ejemplo la administración de la seguridad de la aplicaciones de forma automatizada, ya que dada la cantidad de aplicaciones que hoy están en línea en cualquier centro de cómputos, es muy difícil tenerlo siempre actualizado si se hace manualmente.

En la misma sintonía, resulta fundamental el desarrollo de competencias específicas en ciberseguridad como son las capacidades para testear la seguridad de un sistema así como de poder implementar las mejoras para que un sistema no sea vulnerable. Investigar las nuevas vulnerabilidades, conocer técnicas y herramientas para detectar las mismas, así como estándares para llevar a cabo el proceso de testeo de seguridad, son tareas que forman parte del desafío de contar con sistemas cada vez más seguros en las organizaciones.
La participación en eventos del tipo CTF favorecen el desarrollo de competencias relacionadas a la seguridad de los sistemas, las redes y los servicios, tanto para nuevos recursos humanos que pasan a formar parte del grupo de investigadores como método de capacitación y actualización para aquellos que forman parte del mismo.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

En la actualidad, las principales líneas de trabajo en las que el grupo de investigación en ciberseguridad desarrolla sus actividades incluye:

- Monitoreo de seguridad inteligente.
- Formación y actualización a través del desarrollo de competencias mediante la continua participación en concursos de tipo CTF.
- Detección y análisis de vulnerabilidades en distintos tipos de dispositivos, protocolos y tecnologías.
- Gestión de la seguridad de infraestructura
- Gestión de incidentes de seguridad.
- Forensia Digital.
- Infraestructura de clave pública PKI.
- Desarrollo seguro de software.

3. RESULTADOS OBTENIDOS Y ESPERADOS

Como principales objetivos se plantean:

- Crear un mecanismo alternativo para detectar malware que represente una mejora sobre las herramientas utilizadas actualmente. A su vez éste mecanismo debería aprovechar los resultados del análisis realizado por otras herramientas para aplicar técnicas que permitan tomar la mejor decisión a la hora de determinar si la actividad es maliciosa o no. Dicho mecanismo será integrado al servicio de monitoreo proactivo de seguridad utilizado por CERTUNLP para proteger la red de la UNLP.
- 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 buenas prácticas para 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 distintos proyectos y actividades a los alumnos de las cátedras de grado y postgrado con contenidos afines de nuestra Facultad.

Entre los resultados que se han obtenido en este último tiempo:

- Como parte del proyecto vinculado al

Centro de Excelencia en Ciberseguridad de la ITU[12], del que formamos parte durante el año 2018, se dictaron los cursos "Ciberseguridad: primeros pasos de un gran desafío" y "CSIRT: coordinando prevención, detección, manejo de incidentes, respuesta y mitigación de ciberataques".

- En junio de 2018, los autores de este trabajo formaron parte de la organización junto con la ITU el evento internacional denominado "Ciberseguridad desde el Río de La Plata". El mismo se llevó a cabo en las instalaciones de esta casa de estudios contando con la participación de referentes nacionales e internacionales en diversas temáticas relacionadas con la ciberseguridad. Contó con la participación de más de 150 personas a lo largo de cinco extensas jornadas que involucran disertaciones, workshops y talleres prácticos.
- Desde el año 2017, un grupo de docentes y alumnos trabajan en temáticas relacionadas a las competencias de seguridad, intercambiando experiencias adquiridas en los últimos concursos estudiando nuevas metodologías y nuevas herramientas. Esta actividad que se realiza semanalmente, se enmarca en los grupos de interés definidos por la Secretaría de Innovación de la Facultad de Informática. Entre los logros de esta línea de trabajo se destacan la clasificación a participar en el evento presencial de Cyberex 2018 y la obtención del primer puesto en el CTF llevado a cabo en la conferencia Ekoparty en septiembre de 2018.

4. FORMACIÓN DE RECURSOS HUMANOS

En esta línea de investigación trabaja un grupo de docentes/investigadores del LINTI (Laboratorio de Investigación en Nuevas Tecnologías Informáticas) de la Facultad de Informática de la UNLP (Universidad Nacional de La Plata). Este equipo de trabajo también forma parte de CERTUNLP, el CSIRT Académico de la Universidad Nacional de La Plata [15], ámbito en el cual aplican directamente las temáticas propuestas.

Allí pueden realizarse pruebas y evaluar alternativas en un escenario real; y la experiencia y el desarrollo propuesto enriquece los servicios allí brindados que mejoran la ciberseguridad de la red de la UNLP y de la comunidad en general.

En este marco, se encuentran en desarrollo dos tesis para obtener la Maestría en Redes de Datos: "Detección de ataques de seguridad en redes usando técnicas de ensembling" de la Lic. Paula Venosa y "Optimización del control y administración de la seguridad en una red de servidores. Su implementación como software libre para contribuir con MatFel" del Lic. Einar Lanfranco. También está en curso la tesis del doctorando Ignacio Gallardo Urbini titulada "Estrategia de Ciberseguridad distribuida, aplicando el concepto de operación de inteligencia".

En particular la tesis "Detección de ataques de seguridad en redes usando técnicas de ensembling" sumará también al proyecto SLIPS

---

1 https://academy.itu.int/index.php?option=com_content&view=article&id=154&Itemid=588&lang=en

2 https://academy.itu.int/index.php?option=com_content&view=article&layout=edit&id=249&lang=en

[16] del Stratosphere Laboratory en República Checa, que funciona en el ámbito de la CVUT (Czech Technical University in Prague).

Por otra parte, la tesina de grado de Damián Rubio "Evolución del sistema de gestión de incidentes de seguridad orientado a CSIRT de la UNLP - Ngen" está íntimamente ligada al desarrollo y evolución de NGEN, el sistema de gestión de incidentes actualmente usado en CERTUNLP y liberado como software libre para su uso y contribución por parte de la comunidad.

La tesina en curso "Automatizando la resolución de problemas en competencias de seguridad informática" de los alumnos Jeremías Pretto y Facundo Basso profundiza la investigación y el desarrollo de herramientas utilizables en competencias tipo captura tu bandera.

5. REFERENCIAS

[1] - Díaz, Francisco Javier, Molinari, Lía Hebe, Venosa, Paula, Macia, Nicolás, Lanfranco, Einar Felipe, Sabolansky, Alejandro Javier (2018). WICC 2018 (Workshop de Investigadores en Ciencias de la Computación). UNNE, Corrientes, Argentina. Abril de 2018. Libro de Actas XX Workshop de Investigadores en Ciencias de la Computación. pp1056-1060. ISBN 978-987-3619-27-4.

[2] - CERTunlp: CSIRT académico de la Universidad Nacional de La Plata: http://www.cert.unlp.edu.ar/

[3] - PKIGrid CERTUNLP: https://www.pkigrid.unlp.edu.ar

[4] - Equipo SYPER: https://ctftime.org/team/2003

[5] - Proyecto: F020 - Internet del Futuro. Ciudades Digitales, Inclusivas, Innovadoras y Sustentables, IoT, Ciberseguridad, Espacios de aprendizaje del Futuro. https://cyt.proyectos.unlp.edu.ar/projects/11-f020

[6] - Kasey Panetta. (2017). 5 Trends in Cybersecurity for 2017 and 2018. Diciembre de 2018, de Smarter with Gartner Sitio web: https://www.gartner.com/smarterwithgartner/5-trends-in-cybersecurity-for-2017-and-2018/

[7] - Anna L. Buczak, Member, IEEE, and Erhan Guven, Member, IEEE. (SECOND QUARTER 2016). A Survey of Data Mining and Machine Learning Methods for Cyber Security Intrusion Detection. IEEE COMMUNICATIONS SURVEYS & TUTORIALS, 18, 1153.

Metodología para el análisis de incidentes de ciberseguridad o ciberataques durante las acciones de ciberdefensa de las infraestructuras críticas de la defensa nacional – infoscopia –

Director de Proyecto: Julio C. Liporace[a] jcliporace@est.iue.edu.ar

Adrián Buscaglia[a] abuscaglia@est.iue.edu.ar

Pablo Croci[a] pcroci@est.iue.edu.ar

Nicolás Díaz Pais[a] ndiazpais@est.iue.edu.ar

Darío Fernández[a] dfernandez@est.iue.edu.ar

Verónica Ferreyra[c] vferreyra@fuerzas-armadas.mil.ar

Ignacio Martín Gallardo[ab] igallardo@est.iue.edu.ar

Elvira Quiroga[ab] equiroga@est.iue.edu.ar

Fernando Vera Batista[a] verabatista@est.iue.edu.ar

César D. Cicerchia[a] cdcicerchia@est.iue.edu.ar

a. Facultad de Ingeniería del Ejército (FIE), Ejército Argentino – Universidad de la Defensa Nacional

b. Centro de Investigación y Desarrollo de Sistemas Operacionales (CIDESO), Dirección General de Investigación y Desarrollo (DIGID) - Ejército Argentino

c. Comando Conjunto de Ciberdefensa (CCCD), Estado Mayor Conjunto de las Fuerzas Armadas

RESUMEN

El desarrollo actual de la Ciberseguridad y la Ciberdefensa está apoyado en el empleo de procedimientos reactivos de detección, mitigación y remediación, que se aplican cuando el efecto de la agresión sobre una Infraestructura Crítica del Sistema de Defensa Nacional (ICSDN) ya ocurrió. Se ha adoptado un enfoque derivado de las normativas de Seguridad de la Información, como las formuladas por el estándar internacional ISO/IEC 27000. El aporte original de esta línea de investigación se basa en proponer una nueva metodología de Ciberdefensa, desarrollando un abordaje proactivo hacia las amenazas antes que éstas comprometan la infraestructura crítica, a fin de sorprender al agresor mediante una defensa dinámica y reducir sus posibilidades de éxito.

Las líneas de investigación de InFoscopia tienen como alcance desarrollar una metodología proactiva de análisis de eventos que ocurren antes de que el objetivo cibernético haya sido comprometido. Se enfocará en las ICSDN estratégicas (servicios esenciales de energía, transporte, financieros, comunicaciones e informática, alimentos, agua, químicos, nuclear o espacial) y de capacidades militares.

Contribuirá al desarrollo de procedimientos proactivos por parte de los grupos de respuesta del tipo Computer Security Incident Response Team (CSIRT) o Centro de Operaciones de Ciberdefensa, encargados de la protección de las ICSDN.

Palabras Clave: *Informática Forense. Informatoscopia. Ciberseguridad. Ciberdefensa. Ciberataque. Infraestructura Crítica.*

CONTEXTO

La reciente Directiva de Política de Defensa Nacional [1] destaca que, en la atención del riesgo de "Ataques externos a objetivos estratégicos", el Sistema de Defensa Nacional debe focalizarse en "aquellas infraestructuras cuyo funcionamiento resulte crítico para el cumplimiento de las funciones vitales del Estado Nacional, su Defensa Nacional, el ejercicio de la soberanía y el resguardo de la vida y la libertad de sus habitantes."

El Ejército Argentino (EA), desde hace veinte años patrocina la construcción de su propio sistema de comando y control (SC2) para sus brigadas. Los SC2 son de naturaleza socio – técnica y complejos en su concepción y diseño (Clay, 2007). Por sus características y su finalidad de empleo, constituyen una ICSDN (Dean, 2013). Por otra parte, el EA ha proporcionado los recursos humanos formados como

[1] Decreto 703/2018. DECTO-2018-703-APN-PTE - Directiva de Política de Defensa Nacional. Aprobación. https://www.boletinoficial.gob.ar/pdf/pdfAnexoPrimera/5568234A01.pdf/20180731/0

Ingenieros Militares o en Sistemas de Computación, para la creación y constitución del Comando Conjunto de Ciberdefensa[2].

En base a estos antecedentes, el grupo de trabajo del proyecto InFoscopia es responsable de abordar esta línea de investigación. Para su organización ha recibido el aporte de conocimiento de docentes y graduados de la Especialización en Criptografía y Seguridad Teleinformática y de Ingeniería en Informática, ambas carreras de la misma unidad académica. Asimismo, los docentes son investigadores categorizados con trayectoria en otros proyectos de investigación de la FIE o docentes con experiencia en investigación aplicada al desarrollo tecnológico precompetitivo de sistemas militares.

InFoscopia es un proyecto acreditado por el Programa de Acreditación y Financiamiento de Proyectos UNDEFI, convocado mediante Resolución Rectoral 154/18 de la Universidad de la Defensa Nacional (UNDEF). El financiamiento del proyecto está sustentado por parte de la FIE, mediante la asignación de cargos y horas de investigación de los docentes integrantes del grupo de trabajo, y de la UNDEF, por medio de un subsidio UNDEFI que se renueva anualmente.

Durante las actividades de investigación se intercambiarán conocimientos y experiencias con los siguientes grupos de investigación: Grupo de Investigación del CriptoLAB de la FIE / UNDEF; Grupo de Investigación Sistemas Operativos e Informática Forense de la Facultad de Ingeniería de la Universidad FASTA (sede Mar del Plata).

Por su parte, La Facultad de Ingeniería del Ejército, así como la Facultad de Ingeniería de la Universidad FASTA son miembros fundadores de la Red Universitaria de Informática Forense (UNIF). Esta red constituye un apoyo fundamental para el Grupo de Investigación.

1. INTRODUCCIÓN

Las Infraestructuras Críticas del Sistema de Defensa Nacional (ICSDN) constituyen recursos diversos y complejos, aunque en la actualidad todas tienen uno o más componentes de TIC (Edwards, 2014). Desde la perspectiva del Estado, se debe considerar una infraestructura como aquel conjunto de medios técnicos, servicios e instalaciones necesarios para el desarrollo de las actividades básicas de la sociedad. En este sentido, la mayoría de esas actividades proveen servicios esenciales de carácter estratégico a la sociedad, al gobierno y a los habitantes en su conjunto, tanto sean prestados por organizaciones de gestión pública como privada (Baggett & Simpkins, 2018).

La interrupción o perturbación severa de su funcionamiento, ocasionaría graves efectos sobre el normal desarrollo de las actividades básicas de la sociedad; por tal motivo deben ser consideradas infraestructuras críticas y su defensa, aún en tiempo de paz, es un deber fundamental del Estado. La componente TIC de cualquier ICSDN puede ser afectada desde el Ciberespacio, con un elevado y creciente riesgo que debe ser analizado y prevenido frente a un número importante de amenazas cibernéticas, cuyo comportamiento es dinámico, cambiante y de difícil predicción (Johnson, 2015).

Las amenazas cibernéticas progresan en su acción ofensiva, evolucionando entre las siguientes etapas (The Mitre Corporation, 2018):

- Exploración o Reconocimiento inicial.
- Adquisición intrusiva de servicios o procesos computacionales de la infraestructura / Desarrollo de herramientas acordes a la debilidad a ser explotada.

[2] http://www.fuerzas-armadas.mil.ar/Dependencias-CIBDEF.aspx

- Entrega o distribución
- Compromiso Inicial.
- Uso indebido / Escalamiento de Privilegios.
- Reconocimiento interno.
- Movimiento lateral.
- Establecimiento de la persistencia (consolidación).
- Ejecución de la Misión o cumplimientos de Objetivos.
- Exfiltración.

El conocimiento tecnológico que proporciona la Informática Forense, la Informatoscopia y su apoyo en las Ciencias de la Computación tiene su campo de aplicación en la Justicia (Consejo General del Poder Judicial (CGPJ) et al., 1996). Esta capacidad se sustenta en el empleo de técnicas científicas y analíticas especializadas sobre la infraestructura tecnológica y se desarrolla con la finalidad de identificar, preservar, analizar y presentar datos que sean válidos dentro de un proceso legal, en el marco de un delito informático que es objeto de investigación judicial. Es decir que estas disciplinas actúan *ex post*, luego de ocurrido un delito (Domínguez, 2013).

En la Ciberdefensa tienen prioridad la mitigación, contención o respuesta inmediata para detener los efectos de una ciberagresión o ciberataque, manteniendo el objetivo de la recolección de evidencias de la vulneración de los sistemas y preservar las pruebas, pero dejándolo en segundo orden de prioridad (Intelligence and National Security Alliance (INSA), 2018). Sin embargo, los métodos y herramientas de la Informática Forense pueden resultar de utilidad para apoyar el proceso de gestión de incidentes de Ciberseguridad y Ciberdefensa, pero bajo otra metodología enfocada en una acción proactiva, antes que el efecto de la agresión llegue a su punto culminante (Colbaugh & Glass, 2011).

La explotación del análisis forense digital y su capacidad de extraer muestras o pruebas de las computadoras, equipos móviles y otros dispositivos es fundamental para descubrir e interpretar datos electrónicos.

Un componente de soporte al desarrollo de la metodología es la "línea de tiempo" que permite mostrar quién hizo, qué y cuándo, de manera de poder afirmar de forma concluyente que la Acción A causó el Resultado B (Amusategui López, 2016).

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Se identifican dos líneas principales de investigación de interés para InFoscopia; a saber:

- Análisis de amenazas cibernéticas y de tendencias de ciberagresiones a infraestructuras críticas estratégicas, para modelar el comportamiento de agresores cibernéticos que puedan afectar, especialmente, la disponibilidad[3] de las infraestructuras críticas.
- Métodos de análisis forense digital en memoria, de ingeniería inversa de software y de análisis de malware, que puedan extenderse o aplicarse para estudiar procesos "vivos" en dispositivos de datos, red, seguridad informática o soporte. Se consideran de interés para elaborar un conjunto de Indicadores de Amenaza (*Indicators of Threat* - IoTh), por analogía a los Indicadores de Compromiso (*Indicators of Compromise* - IoC).

Para este proyecto, la primera línea de investigación resulta de interés para el modelado del comportamiento del agresor cibernético, especialmente en las etapas previas del ciberataque

[3] *Availability*, adoptada por el estándar internacional ISO/IEC 27000. Se refiere a la "propiedad de la información [o de un sistema] de estar accesible y utilizable cuando lo requiera una entidad autorizada".

(exploración o reconocimiento inicial, adquisición intrusiva de servicios o procesos computacionales de la infraestructura, desarrollo de herramientas ofensivas acordes a la debilidad a ser explotada) y en momento inicial del asalto a la infraestructura, cuando empieza la entrega o distribución de una ciberarma (malware, tramas anómalas de datos, etc).

La segunda línea de investigación tiene interés, entre otros puntos, sobre los métodos de recolección de información disponibles para archivos dependientes de la memoria RAM, como ser la RAM propiamente dicha y el archivo de paginación de los sistemas operativos. También, el análisis de artefactos o piezas de software en tiempo de ejecución, los procesos desatados por el malware o software dañino y el análisis de tráfico en la red, desde y hacia el activo comprometido. Se pretende analizar y comprender eventos en tiempo de ejecución para clasificarlos como normales o anómalos con un grado de confianza aceptable.

El proyecto tiene previsto tres etapas:

- Formulación metodológica para activos bajo entornos Windows en redes Ethernet proponiendo indicadores de amenazas a la disponibilidad. Se validará con pruebas de concepto.
- Ampliación de la metodología para ambientes heterogéneos con dispositivos Linux y medios de redes y seguridad particulares de infraestructuras críticas. Se desarrollará un prototipo experimental.
- Se completará con la experimentación y selección de parámetros definitivos de los indicadores de amenazas definidos.

3. RESULTADOS OBTENIDOS/ESPERADOS

El principal resultado esperado es desarrollar una metodología propia de análisis de eventos o incidentes aplicable a las fases iniciales de una ciberagresión o ciberataque que pudiera ocurrir en uno o más activos esenciales de una ICSDN, con la finalidad de dar respuesta a los siguientes propósitos:

- Detectar cómo y cuándo ocurrió una violación de la protección de una ICSDN.
- Identificar los activos esenciales o sistemas comprometidos y afectados.
- Determinar qué atacantes tomaron o cambiaron procesos.
- Contener y remediar los incidentes que se configuren.
- Desarrollar indicadores y fuentes clave de inteligencia de amenazas.
- Buscar violaciones adicionales usando el conocimiento de las tácticas, técnicas y procedimientos del agresor.

Con respecto a la primera línea, se ha comenzado por adoptar como referencia el marco de trabajo MITRE's Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK), desarrollada y mantenida por la organización estadounidense The Mitre Corporation. Se trata de una base de conocimiento accesible a nivel mundial, sobre tácticas, técnicas y procedimientos que utilizan los agresores cibernéticos.

Entre los primeros resultados de la segunda línea de investigación, por medio del empleo de varias herramientas forenses de análisis de memoria de computadoras con arquitectura *Von Neumann*, se logró validar la técnica de obtención de datos residentes en memoria para el análisis de cadenas de caracteres presentes en tiempo de ejecución. Como continuación de este trabajo, mediante la utilización de la herramienta *Sysmon* (de la suite *Sysinternals* de Microsoft) se pretende estudiar el comportamiento de los activos bajo condiciones normales de operación y compararlos con los efectos de eventos asociados a la fase previa de los ciberataques. La finalidad es obtener información detallada sobre las creaciones de procesos, conexiones de red y cambios en el tiempo de creación de archivos. Al recopilar los eventos que se generan, se podrá realizar el análisis en vivo de los *Logs* para identificar actividades anómalas y comprender cómo operan los intru-

sos y el *malware* en la red de una infraestructura crítica, de manera de encontrar posibles Indicadores de Amenazas (IoTh) en un momento previo al compromiso de la disponibilidad del activo (*weaponization* en el *framework* PRE-ATTACK de Mitre).

4. FORMACIÓN DE RECURSOS HUMANOS

El Grupo de Investigación tiene conocimientos y experiencia sobre Informática Forense, Ciberseguridad, Ciberdefensa, Redes de Información, Sistemas Operativos, Sistemas de Control, Sistemas Electrónicos, Ingeniería de Software y Programación.

La estructura del equipo de trabajo está conformada por dos docentes e investigadores del posgrado de Especialización en Criptografía y Seguridad Teleinformática, tres docentes e investigadores de Ingeniería en Informática. Además, un tecnólogo posgraduado en la especialización mencionada, con dos tesis de posgrado aprobadas, vinculadas al proyecto; actualmente, realiza el Doctorado en Ciencias Informáticas con tesis afín al proyecto. El grupo se completa con dos tecnólogos especializados en sistemas operativos y redes de computadoras, y un tercer tecnólogo con experiencia en incidentes de ciberseguridad. A partir de este año, se han incorporado dos alumnos de la carrera de Ingeniería en Informática y uno de Ingeniería Electrónica que desarrollarán tesinas de grado vinculadas al proyecto.

La formación de recursos humanos es un objetivo del proyecto y está reflejado en la constitución del grupo, donde se integran los conocimientos de docentes muy experimentados y jóvenes, alumnos de ingeniería, personal técnico y tecnólogos de tres laboratorios asociados al proyecto (CriptoLab y CIDESO / Ejército; InFo-Lab / Universidad FASTA). Por último, se previeron otorgar becas de doctorado, postgrado y grado e incorporar pasantes en el marco de la Red UNIF.

5. BIBLIOGRAFÍA

Amusategui López, J. M. (22 de Dic de 2016). *Universitat Oberta de Catalunya.* Obtenido de http://openaccess.uoc.edu/webapps/o2/bitstream/10609/60765/6/jamusateguiTFM1216memoria.pdf

Baggett, R. K., & Simpkins, B. K. (2018). *Homeland Security and Critical Infrastructure Protection, 2nd Edition.* Santa Barbara, California. USA: Praeger Security International.

Clay, W. (2007). *Network Centric Operations: Background and Oversight Issues for Congress.* CRS Report for Congress.

Colbaugh, R., & Glass, K. (15 de Ago de 2011). *IEEE Xplore Digital Library.* (P. o. Informatics, Ed.) doi:10.1109/ISI.2011.5984062

Consejo General del Poder Judicial (CGPJ) et al. (31 de Dic de 1996). *CITA.ES.* Obtenido de http://cita.es/apedanica/INFORMAT.HTM

Dean, S. E. (2013). *Cyber Defense: Securing Military Systems and Critical Civilian Infrastructure from an Electronic 9/11.* Hampton Roads International Security Quarterly.

Domínguez, F. L. (2013). *Introducción a la Informática Forense.* Madrid. España: Ra-Ma.

Edwards, M. (2014). *Critical Infraestructurre Protection.* Ankara, Turquía: IOS Pres BV.

Intelligence and National Security Alliance (INSA). (31 de Ago de 2018). *Intelligence and National Security Alliance.* Obtenido de https://www.insaonline.org/wp-content/uploads/2018/08/INSA-Managing-Cyber-Attack-Critical-Infrastructure.pdf

Johnson, T. A. (2015). *Cybersecurity: Protecting Critical Infrastructures from Cyber Attack and Cyber Warfare.* St.Louis, Missouri. USA: CRC Press.

The Mitre Corporation. (22 de Nov de 2018). *MITRE ATT&CK*. Obtenido de https://attack.mitre.org/

Métodos, técnicas y estrategias para mejorar la seguridad de los sistemas de software.

Agustín Ferrari, Edgardo Bernardis, Mario Berón, Hernán Bernardis,
Maria Joao Tinoco Varanda Pereira, Daniel Riesco

Departamento de Informática
Universidad Nacional de San Luis
Ejército de los Andes 950, San Luis, Argentina
Departamento de Informática e Comunicações
Instituto Politécnico de Bragança
Bragança, Portugal

{ebernardis, mberon, hbernardis, driesco}@unsl.edu.ar,
ferrariagustin93@gmail.com, mjoao@ipb.pt

RESUMEN

Con el crecimiento de internet y las distintas dinámicas de la sociedad actual; ha cambiado en gran medida la forma de interactuar e intercambiar información entre las personas y las empresas. Este intercambio se vuelve blanco de ataques por parte de todos aquellos actores que quieren obtener información útil y valiosa a sus propios intereses o de terceros. Ante este panorama se vuelve imperioso implementar todo tipo de medidas y acciones tendientes a evitar estos ataques, por tal motivo nace lo que se denomina Seguridad Informática. Toda acción, herramienta o metodología enfocada a evitar, contrarrestar o retrasar ataques contra activos sensibles juega un rol sumamente importante para los diversos actores.

Por lo antes descrito, en este artículo se presenta una línea de investigación que aborda el diseño y construcción de una herramienta cuyo principal objetivo es desarrollar e incrementar el nivel de seguridad de Servicios Web.

Palabras claves: Seguridad Informática, activos sensibles, Servicios Web.

CONTEXTO

La presente línea de investigación se enmarca en el proyecto (P-031516.) 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, y el proyecto (PO/16/93.) denominado: "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.

1. INTRODUCCIÓN

Actualmente se están popularizando los Web Services como artefactos de software a partir de los cuales se pueden construir sistemas más complejos. Según la W3C, un Web Service es: "Una aplicación de software identificada por una URI, cuya interfaz y enlaces son capaces de ser definidos, descritos 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, algunos de ellos se publican al resto del mundo de manera libre, mientras que otros son utilizados de manera interna por sus equipos de

desarrollo. Esto permite que cada equipo de desarrollo elija la arquitectura que desee para construir sus proyectos sin afectar la vinculación con el resto del sistema y/o proyectos. 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.

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 para la definición de los mismos.

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) [6]. 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 [7]. Todos contienen información formal (código fuente) e informal (identificadores, comentarios, documentación, etc.) y es en este tipo de información en donde los atacantes hacen foco para obtener información beneficiosa a sus propósitos. Suena lógico entonces incrementar la seguridad que posee un determinado WSDL para evitar e impedir los ataques.

La línea de investigación comprende los principales lineamientos de:

- *Seguridad Informática*: Se estudia el concepto de seguridad informática y sus principios generales.
- *Ofuscación*: Se estudia la ofuscación de código y sus diferentes clases.
- *Reflexión*: Se estudia el patrón de reflexión a fin de auto-incorporar algoritmos de seguridad en ejecución.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

En las siguientes secciones se describen las líneas de investigación que llevan adelante los autores de este artículo.

2.1. Seguridad Informática

El objetivo de la Seguridad Informática es obtener un nivel aceptable de seguridad, entendiéndose por aceptable un nivel de protección suficiente para que la mayor parte de potenciales intrusos, interesados en los equipos con información de una organización o persona, fracasen en cualquier intento de ataque contra los mismos. Asimismo, se encarga de establecer los mecanismos para registrar cualquier evento fuera del comportamiento normal y tomar las medidas necesarias para re-

establecer las operaciones críticas a la normalidad [1].

Los principios generales de la seguridad de la información [2, 3] son:

- Integridad: implica que debe salvaguardarse la totalidad y la exactitud de la información que se gestiona. A su vez, puede incluir:
  - ❑ Autenticidad: autoría indiscutible. Definir que la información requerida es válida y utilizable en tiempo, forma y distribución.
  - ❑ No Repudio: la seguridad de que una parte no puede negar posteriormente los datos de origen; suministrando la prueba de integridad y el origen de los datos, y que puede ser verificado por un tercero. Evitar que cualquier entidad que envió o recibió información alegue, ante terceros, que no la envió o recibió.
- Confidencialidad: implica que debe protegerse la información de forma tal que sólo sea conocida y accedida por las personas autorizadas y se la resguarde del acceso de terceros.
- Disponibilidad: implica que debe protegerse la información de forma tal que se pueda disponer de ella para su gestión en el tiempo y la forma requerida por el usuario.

2.2. Ofuscación

En términos generales, se entiende por ofuscar un código fuente o un código intermedio, un proceso mediante el cual se transforma mediante la aplicación de algoritmos de reescritura, un código perfectamente legible y entendible por una persona en otro de funcionalidad equivalente en un ciento por ciento, pero, en términos ideales, totalmente ilegible e incomprensible para un lector.

En pocas palabras, algunas de las técnicas de ofuscación más comunes consisten en la inclusión de bucles irrelevantes, cálculos innecesarios, comprobaciones fuera de contexto, nombres de funciones y de variables que no tienen nada que ver con su cometido, funciones que no sirven para nada, interacciones inverosímiles entre variables y funciones, etc. Otras técnicas, sin embargo, son mucho más potentes, en el sentido de que requieren un conocimiento superior de las características del lenguaje, e incluso pueden estar diseñadas para burlar a herramientas de ingeniería inversa específicas [10].

La ofuscación de código aplica una o más transformaciones de código que hacen que el código sea más resistente al análisis y la manipulación, pero preservan su funcionalidad [4].

2.2.1. Transformaciones de Código.

Las transformaciones de código para ofuscar un programa se pueden dividir en cuatro clases principales [5], estas son:

- Transformaciones Léxicas o de Diseño: afectan la información en el código que es innecesaria para su ejecución y que reduce la información disponible para un lector humano. Ejemplo de esto es la codificación de nombres, la eliminación de comentarios, etc.
- Transformaciones de Flujo de Control: actúan modificando el flujo de control del programa. Por ejemplo inserción de predicados opacos[1], que agregan transferencias de control extra que nunca se tomarán en tiempo de ejecución.
- Transformaciones de Flujo de Datos: operan sobre las estructuras de datos usadas en el programa. Un ejemplo de una transformación de datos es desmantelar arreglos para aumentar la complejidad del código.
- Transformaciones Preventivas: intentan detener el funcionamiento correcto de los decompiladores o desofuscadores de código. Ejemplo de esto es insertar "bytes basura" entre otras instrucciones. Un algoritmo de desensamblaje interpreta estos bytes como el comienzo de una instrucción, y se arriesga a desmontar incorrectamente los bytes de la instrucción original.

[1] Informalmente, una variable V (o predicado P) es opaca si tiene alguna propiedad q que conoce a priori el ofuscador, pero que es difícil de deducir para el desofuscador [12].

2.3 Reflexión

Durante la ejecución de una aplicación el código de la misma permanece estático; las instrucciones de ejecución no cambian. Se dice que una aplicación contiene código automodificable cuando es capaz de modificar sus propias instrucciones de código, además de sus datos, durante su ejecución [11]. Al contener esta característica y hacer uso de ella de forma correcta, se pueden lograr grandes resultados como modificar funcionalidades o incorporar herramientas a un sistema en ejecución.

En el caso de esta investigación, se propone que el sistema permita al usuario incorporar algoritmos de ofuscación en ejecución y así poder aumentar la seguridad de los servicios web, es decir, incorporar un módulo donde el usuario pueda desarrollar sus propios algoritmos que luego serán utilizados para la protección de los Servicios Web. Para llevar a cabo esta funcionalidad se hace uso de la Reflexión.

La Reflexión es: *la capacidad integral de un programa para observar o cambiar su propio código, así como todos los aspectos de su lenguaje de programación (sintaxis, semántica o implementación), incluso en tiempo de ejecución. Se dice que un lenguaje de programación es reflexivo cuando proporciona a sus programas la capacidad de reflexión* [8]. La palabra integral es muy importante, la verdadera reflexión no impone límites a lo que un programa puede observar o modificar. En el caso de esta investigación, se hace uso de esta capacidad a fin de permitirle al usuario incorporar algoritmos personalizados que aumenten la seguridad de los servicios web.

3. RESULTADOS OBTENIDOS/ESPERADOS

El trabajo de investigación de esta línea permitió obtener los siguientes resultados:

- Se desarrolló ATENOS una herramienta que permite recorrer Servicios Web, extraer sus identificadores, protegerlos y generar un nuevo servicio con funcionalidad equivalente pero más seguro[2].
- Se logró mejorar significativamente la seguridad del WSDL al ofuscar los tags del Servicio Web dificultando su entendimiento y por consiguiente disminuyendo considerablemente el tipo de ataques que se pueden realizar sobre el mismo.
- Se desarrollaron e implementaron tres algoritmos para ofuscar, los tags de un WSDL, mediante la aplicación de diferentes técnicas de transformaciones léxicas sobre la información tomada como objetivo para incrementar su seguridad.
- Se logró crear una plataforma que permite agregar modificar y eliminar algoritmos propios de los usuarios; siendo la misma un rasgo distinguible del resto de las herramientas de similares características.
- Se permite utilizar también la encriptación de los archivos al brindar la opción de desencriptar el archivo.
- Se utilizó ATENOS con un ejemplo de aplicación de la vida real el caso de la Facturación Electrónica perteneciente a la AFIP [9].

4. FORMACIÓN 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.

---

[2] Cuando se menciona la extracción de identificadores, se hace referencia al proceso de construcción del árbol de sintaxis abstracta (AST) y la aplicación de diferentes recorridos para recuperar/modificar la información.

5. BIBLIOGRAFÍA

[1] Alejandra Stolk. Técnicas de seguridad informática con software libre, 2013. Parque Tecnológico de Mérida. ESLARED.

[2] Salton, International Telecommunication Union. SERIES X: DATA NETWORKS, OPEN SYSTEM COMMUNICATIONS AND SECURITY. Telecommunication security. Overview of cybersecurity. 2008.

[3] Salton, Borghello, Cristian F. Seguridad Informática, sus implicancias e implementación. Tesis Licenciatura en Sistemas. Universidad Tecnológica Nacional. 2001.

[4] Code obfuscation techniques for software protection. Cappaert, Jan. Katholieke Universiteit Leuven. 2012.

[5] A taxonomy of obfuscating transformations. Collberg, Christian and Thomborson, Clark and Low, Douglas. 1997.

[6] WSDL Specification for W3C https://www.w3.org/TR/wsdl..

[7] 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.

[8] Malenfant, J., Jacques, M., & Demers, F. N. (1996, April). A tutorial on behavioral reflection and its implementation. In Proceedings of the Reflection (Vol. 96, pp. 1-20).

[9]AFIP.https://wswhomo.afip.gov.ar/wsfev1/service.asmx. 2018.

[10] DOLZ, Daniel; PARRA, Gerardo. Ofuscadores de código intermedio. En *VIII Workshop de Investigadores en Ciencias de la Computación*. 2006.

[11] XIANYA, Mi, et al. A survey of software protection methods based on self-modifying code. En *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*. IEEE, 2015. p. 589-593.

[12] Christian Collberg, Clark Thomborson, and Douglas Low. Manufacturing cheap, resilient, and stealthy opaque constructs. ACM, New York, NY, USA, 184-196. 1998.

Modelado Probabilístico Basado en Aprendizaje Profundo para la Detección de Anomalías en el Tráfico de Red

Santiago Eguren[1], Carlos A. Catania[12], Jorge Guerra[2]

[1]Instituto de Investigaciones, Facultad de Informática y Diseño, U. Champagnat, Belgrano 721, 5501 Godoy Cruz, Mendoza, Argentina.

[2]Facultad de Ingeniería, Universidad Nacional de Cuyo
Campus Universitario, Mendoza, Argentina.

santiago.e@live.com.ar, harpo@ingenieria.uncuyo.edu.ar, jguerra@uncu.edu.ar

RESUMEN

En la detección de intrusos en el tráfico de red pueden distinguirse dos estrategias: la detección por mal uso y la detección por anomalías. En la primera, se parte de patrones conocidos que representan el comportamiento malicioso y posteriormente se identifica a los mismos en el tráfico de red. Mientras que en la segunda, se modela el tráfico normal y luego se asume que toda instancia de tráfico que no se adapte al modelo se trata de un comportamiento malicioso. Uno de los mayores inconvenientes de la detección por mal uso es no es capaz de reconocer tipos de ataques que no se encuentran en el conjunto de patrones de ataque previamente conocidos. Por su parte la ventaja de la detección de anomalías es que son capaces de detectar nuevos ataques, sin embargo puede en muchos casos clasificar como maliciosas a instancias de tráfico normal. El presente proyecto se centra en la generación de modelos probabilísticos para la detección de anomalías en el tráfico de red. En particular, se propone analizar la utilización de un enfoque de aprendizaje de máquinas basado en redes neuronales profundas, las cuales han demostrado una gran eficacia en la detección de patrones en diversas áreas.

Palabras Claves: Seguridad de Redes, Detección de Anomalías, Aprendizaje Automático

CONTEXTO

El presente proyecto 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. Este trabajo es parte del proyecto de investigación que dio inicio en diciembre de 2017 en el marco de los proyectos bienales de investigación de la Facultad de Informática y Diseño de la Universidad Champagnat. Es importante destacar que el proyecto cuenta con la colaboración de investigadores de la facultad de ingeniería de la Universidad Nacional de Cuyo.

1 INTRODUCCIÓN

En los últimos años garantizar la seguridad de los datos en las redes de computadoras se ha transformado en un serio problema. En la actualidad un especialista en seguridad de redes debe estar alerta para detectar posibles ataques, informándose de las nuevas vulnerabilidades descubiertas o tipos de ataques perpetrados. Para dar soporte a esta

tarea se cuenta con herramientas para el análisis y detección de intrusos (IDS).

En la detección de intrusos en el tráfico de red se distinguen dos estrategias: la detección por mal uso y la detección por anomalías [1]. En la primera, se parte patrones conocidos que representan el comportamiento malicioso y posteriormente se trata de identificar a los mismos en el tráfico de red. Mientras que en la segunda, se modela el tráfico normal de red y luego se asume que toda instancia de tráfico que no se adapte al modelo puede deberse a un comportamiento malicioso.

La mayoría de los sistemas de detección de intrusos actuales se basan en detección por mal uso. Uno de los mayores inconvenientes es que no son capaces para reconocer tipos de ataques que no se encuentran en el conjunto de patrones de ataque previamente conocidos. Por su parte la ventaja de la detección de anomalías es que puede ser de gran ayuda en la detección de nuevos ataques, sin embargo presenta el inconveniente de que en muchos casos clasifica como maliciosas a instancias de tráfico normal. Lo que se conoce como falsos positivos.

De detallan a continuación los enfoques más relevantes aplicados en los últimos años al problema de la detección de intrusos en el tráfico de red [2].

Modelos basados en reglas: En este enfoque cada posible incidente de seguridad es descrito mediante una regla o firma utilizando un lenguaje específico. Posteriormente, cuando se observa en el tráfico de red alguna regla previamente especificada, se dispara una alarma. Este tipo de sistemas son los más frecuentemente utilizados en la actualidad. SNORT [3], quizás el sistema de detección más utilizado en la actualidad, sigue una estrategia de detección por mal uso aplicando un enfoque basado en reconocimiento de de reglas.

Modelos Probabilísticos: La idea detrás de estos métodos consiste en el análisis del tráfico de red y la construcción de un modelo que describa su comportamiento. La construcción del modelo normalmente se basa en ciertas métricas obtenidas a través del tiempo [6,7,8,9]. Un enfoque común para la generación de estos modelos consiste en la aplicación de técnicas de Aprendizaje de Máquinas. Dichas técnicas permiten la construcción automática de los modelos de la red a partir de lo datos previamente recopilados. Entre las muchas técnicas de aprendizaje de máquina aplicadas al problema de detección de intrusos se destaca [8,9,10,11,12,13].

El presente proyecto se centra en la generación de modelos probabilísticos para la detección de anomalías en el tráfico de red. En particular, se propone analizar la utilización de un enfoque de aprendizaje de máquinas basado en redes neuronales profundas. En particular se propone evaluar la aplicación de las redes recurrentes de tipo Long Short Term Support (LSTM) [14], las cuales han resultado adecuadas para en análisis de secuencias. Se considerarán además, las redes convolucionales multi-capa (ConvNets), las cuales han demostrado una gran eficacia en la detección de patrones en diversas áreas.

2 LÍNEAS DE INVESTIGACIÓN Y DESARROLLO.

El proyecto se enmarca en dos de las áreas de investigación del instituto de investigaciones de la universidad, en particular la Captura y procesamiento de datos a gran escala y el aprendizaje estadístico.

Para el desarrollo del presente proyecto pueden diferenciarse 3 etapas principales:

A) Análisis preliminar del problema. Las tareas asociadas a esta etapa tienen por objetivo conocimiento de los problemas asociados a la detección de anomalías como así también los modelos probabilísticos actualmente implementados. Una tarea fundamental consistirá en realización de una breve revisión de la literatura sobre la

aplicación de las técnicas de aprendizaje profundo a problemas de detección de comportamiento malicioso. Dicha tarea tiene por objetivo el de tratar de determinar cuáles han sido los beneficios e inconvenientes al aplicar este tipo de algoritmos. Finalmente como última tarea de esta primer etapa se realizará un análisis estadístico descriptivo de diversos conjuntos de datos provistos por el proyecto Stratosphere IPS a fin de analizar diversos aspectos del tráfico de red.

B) Desarrollo de un algoritmo para el reconocimiento de anomalías basado en técnicas de aprendizaje profundo. Esta etapa tiene por objetivo el desarrollo, evaluación y puesta a punto de un primer prototipo funcional para la detección de anomalías en el tráfico de red. Este primer prototipo será desarrollado utilizando alguna de las bibliotecas disponibles que permitan la implementación de modelos basados en aprendizaje profundo de manera simple y eficiente. En esta etapa se definirán los diferentes aspectos de la red como ser: el tipo de red a utilizar, la topología de la red y la secuencia de entrada entre otras. Luego se evaluarán los resultados utilizando un conjunto de datos conteniendo tráfico etiquetado (malicioso y normal).

C) Experimentación. Finalmente en la última etapa se centrará en la evaluación del algoritmo propuesto sobre distintos conjuntos de datos: En particular, se evaluará el algoritmo propuesto con diferentes conjuntos de datos previamente etiquetados. Durante este proceso se consideran las métricas habituales en el área utilizando mecanismos para validar la generalidad del modelo obtenido como validación cruzada. Durante esta etapa se realizaran también estudios comparativos con otros modelos de reconocimiento de anomalías de la literatura.

3 RESULTADOS ESPERADOS

Al término de los dos años de duración del plan de trabajo se pretende que se haya logrado:

1. Fortalecer la línea de investigación en la aplicación de modelos probabilísticos relacionados con el tráfico de red.
2. Obtener una implementación funcional del modelo probabilístico para la detección de anomalías basado en redes neuronales con aprendizaje profundo.
3. Incrementar la experiencia para la posterior aplicación de modelos probabilísticos basados en aprendizaje profundo a nuevas líneas de investigación relacionadas con problemas de ciencia y tecnología.

4 FORMACIÓN DE RECURSOS HUMANOS

Se espera capacitar en el ámbito de la investigación a profesores y alumnos interesados en participar en un entorno académico y tecnológico innovador y a todos aquellos actores interesados en los resultados del proyecto.

Sobre la temática de este proyecto se está trabajando en:

• La tesis doctoral de Jorge Guerra, en el doctorado en Ciencias Informáticas de la Universidad Nacional del centro de la provincia de Buenos Aires.

5 BIBLIOGRAFÍA

[1] Mukherjee, B., Heberlein, L., & Levitt, K. (1994). Network intrusion detection. Network, IEEE,8, 26 –41.
[2] Sperotto, A.; Schaffrath, G.; Sadre, R.; Morariu, C.; Pras, A. & Stiller, B. An Overview of IP Flow-Based Intrusion Detection Communications Surveys Tutorials, IEEE, 2010, 12, 343 -356
[3] Wu, S. X. & Banzhaf, W. The use of

computational intelligence in intrusion detection systems: A review Applied Soft Computing, 2010, 10, 1 - 35
[4] Holland, J. Adaptation In Natural and Artificial Systems The University of Michigan Press, 1975
[5] Goldberg, D. Genetic Algorithms in search Optimization and Machine Learning. Addison Wesley,1989
[6] Lakhina, A.; Crovella, M. & Diot, C. Diagnosing network-wide traffic anomalies SIGCOMM Comput. Commun. Rev., ACM, 2004, 34, 219-230
[7] Stoecklin, M. P.; Le Boudec, J.-Y. & Kind, A. A two-layered anomaly detection technique based on multi-modal flow behavior models Proceedings of the 9th international conference on Passive and active network measurement, Springer-Verlag, 2008, 212-221
[8] Mahoney, M. & Chan, P. Learning rules for anomaly detection of hostile network traffic Data Mining, 2003.ICDM 2003. Third IEEE International Conference on, 2003, 601 - 604
[9] Livadas, C.; Walsh, R.; Lapsley, D. & Strayer, W. Using Machine Learning Techniques to Identify Botnet Traffic Local Computer Networks, Proceedings 2006 31st IEEE Conference on, 2006, 967 -974
[10] C. Catania and C. G. Garino, "Una propuesta de reconocimiento de patrones en el tráfico de red basada en algoritmos genéticos," in ninth Argentinian Symposium on Artificial Intelligence, ASAI, D. Godoy and A. Maguitman, Eds., 2007, pp. 174–185.
[11] Catania, C.; Bromberg, F. & Garcia Garino, C. Bromberg, F. & Verdun, L. (Eds.) An autonomous labeling approach to SVM algorithms for network traffic anomaly detection Proceedings of ASAI 2009 Argentine Symposium on Artificial Intelligence, 2009, 144-15. ISSN 1850-2784.
[12] Catania C., Bromberg F. y García Garino C. "Detección de intrusos en el tráfico de red Mediante Máquinas de Vectores Soporte". En los anales de V EnIDI. Encuentro de Investigadores y Docentes de Ingeniería. San Rafael, 2009. pp. 168-182. ISBN 978-950-42-0087-1.
[13] Catania C., Bromberg F. y García Garino C. "An Autonomous Labeling approach to Support Vector Machines Algorithms for Network Traffic Anomaly Detection". Expert Systems with Applications. Elsevier.DOI:10.1016/j.eswa.2011.08.068
[14] Pablo Torres, Carlos Catania, Sebastian Garcia, and Carlos Garcia Garino "An Analysis of Recurrent Neural Networks for Botnet Behavior Detection". Congreso Bienal de IEEE Argentina (ArgenCon), 2016. Buenos Aires. Argentina.

Tecnología Informática

Aplicada en Educación

Abordaje de expresiones regulares y sus derivadas mediante la utilización del software JFLAP

Elisa Silvia Oliva[1], Mathias G. Diaz Ogas[1], Ana Laura Molina[1], Nancy E. Alonso[1]

[1]Departamento de Informática, Facultad de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de San Juan

eoliva@iinfo.unsj.edu.ar, mathiasgdiaz@gmail.com, lauramolina@outlook.com, nalonso@unsj-cuim.edu.ar

RESUMEN

Realizar actualización de contenidos e introducir uso de Tecnologías de la Información y la Comunicación (TIC) en asignaturas, es una tarea que impacta: tanto en la calidad de la educación, como en la curiosidad y motivación del alumno, al incluir software en los procesos de aprendizaje.

Es proyecto de Matemática Discreta con el uso de JFLAP1: herramienta que permite visualizar gráficamente la construcción de las máquinas de estados finitos; haciendo que el estudiante pueda comprender el funcionamiento de los modelos teóricos incluidos en la asignatura y también, posibilita la autocorrección de ejercicios.

La propuesta se sustenta en adicionar al software desarrollos teórico-prácticos que permitan desde las reglas de la gramática del Autómata Finito Determinístico (AFD) y su Expresión Regular (ER), obtener las reglas de la gramática del AFD que corresponde a la derivada de la ER, lo cual permitiría el trabajo de la derivada de la ER con el software, pues no lo aborda entre sus tareas realizables.

El desarrollo de algoritmos que generen la gramática del AFD de la derivada de una ER, permitirán la interrelación con JFLAP, potenciando así sus capacidades para dar respuesta a la representación gráfica del AFD de la derivada de una ER.

Palabras clave: JFLAP; Autómata Finito Determinístico; Derivada de una Expresión Regular; Algoritmo; TIC.

CONTEXTO

La línea de investigación descripta en este artículo se desarrolla en un proyecto de investigación de cátedra de la asignatura Matemática Discreta en las áreas de autómatas y lenguajes formales. Se realizan desarrollos teóricos y se implementa tecnología aplicada al aprendizaje. Explorando las características de ambientes de aprendizaje en ciencias de la computación enriquecidos con TIC.

La asignatura mencionada pertenece a las carreras de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información, del Departamento de Informática de la Facultad de Ciencias Exactas, Físicas y Naturales (FCEFyN) de la Universidad Nacional de San Juan (UNSJ).

1. INTRODUCCIÓN

En los estándares de competencias sobre el manejo de 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". La habilidad de "aprender a aprender" [1], [2], socialmente requiere la

[1] JFLAP es un software para experimentar con temas de lenguajes formales, incluyendo autómatas finitos no deterministas, autómatas no deterministas, Máquinas Turing de cintas múltiples, varios tipos de gramáticas, análisis y sistemas L.
Disponible en: http://www.jflap.org/

formación de personas capaces de realizar un manejo autónomo de herramientas cognitivas. Estos conceptos son propios de una educación centrada en quien aprende, que es transversal para actuar en cualquier área del mundo laboral o en la continuidad de estudios de nivel superior.

En este contexto los estudiantes imponen un fuerte desafío a las instituciones educativas, el rol de las mismas ya no está en asegurar la estandarización de resultados generando como producto final un alumno "educado", sino más bien en generar oportunidades en la persona para el desarrollo de habilidades y competencias que colaboren en su formación permanente y en lograr que el alumno pueda desarrollar en su camino la autoevaluación crítica y el conocimiento de sí mismo, que pueda conocer cómo aprende y qué necesita para contribuir en un mundo cuyas estructuras sociales, económicas y culturales están en permanente cambio.

Este escenario actual impone acciones que deberán, a través de la integración de las TIC, cubrir las competencias digitales que los alumnos deben incorporar en su bagaje cultural y de conocimientos para desarrollar su autonomía en el camino de aprender a aprender.

La aplicación de metodologías activas para que el alumno logre una autoevaluación permanente es una meta a alcanzar en todos los ciclos de la Educación Universitaria y en asignaturas de índole formal y teórico que abarcan Lenguajes, Computación y Sistemas Inteligentes.

Avances tecnológicos-educativos de las últimas décadas y el proceso de acreditación de carreras de grado han producido cambios curriculares de las carreras de Licenciatura en Ciencias de la Computación y en Sistemas en Información, apoyando el desarrollo de muchos temas de aplicación y de fundamentación matemática, entre los que se encuentran los métodos discretos.

La informática teórica, se sustenta en una base matemática derivada del Álgebra y con fuertes bases de Lógica Matemática. La asignatura Matemática Discreta tiene entre sus objetivos proporcionar los fundamentos matemáticos para la teoría de autómatas y lenguajes formales, como los autómatas finitos: determinísticos y no determinísticos, las expresiones regulares, las derivadas de expresiones regulares, las gramáticas incontextuales, etc. Estos temas conforman los cimientos de Teoría de la Computación.

Estos fuertes pilares matemáticos hacen que se corra riesgo de rechazo por parte del alumno a la hora de aplicar los conocimientos teóricos en los contenidos prácticos. Para salvar este problema y dado que los alumnos no son matemáticos expertos, se busca inicializarlos en conocer los cimientos fundacionales de la informática (lenguajes formales, gramáticas, autómatas, etc.), subrayando los aspectos esenciales de la disciplina que permanecen inalterables ante el cambio tecnológico. Por ello, la metodología aplicada, está enfocada en que los alumnos aprendan a desarrollar su capacidad lógico-matemática, lograr una participación activa en cada uno de los procesos de enseñanza-aprendizaje y, con el uso de nuevas tecnologías, lograr una síntesis entre la teoría y la práctica.

Meirieu [3] sugiere entusiasmar a los alumnos: para lo cual deben emplearse estrategias "accesibles y difíciles al mismo tiempo". Estas estrategias son tales cuando el alumno siente que es capaz de lograr el objetivo propuesto y percibe la existencia de una hipótesis que todavía no le es propia. Por su parte, Maggio [4] distingue dos modalidades al incorporar TIC a las prácticas de enseñanza, inclusiones efectivas e inclusiones genuinas. La inclusión genuina, se refiere a un proceso de integración de TIC de orden epistemológico que "reconoce el complejo entramado de la tecnología en la construcción del conocimiento en modos específicos por campo disciplinar y emula ese entramado en el plano de la práctica de la enseñanza". De esta forma, en lugar de ocupar el lugar subsidiario de la superficie y el agregado, los desarrollos tecnológicos pasan a formar parte del cuerpo mismo del área en la que han sido incluidos.

En la asignatura se hace uso del software libre JFLAP, con el objeto de que el alumno pueda comprender el funcionamiento de los

modelos teóricos incluidos en la asignatura. Esta herramienta permite visualizar simbólica y gráficamente la construcción y el comportamiento de las máquinas de estados finitos.

Una de las utilidades de JFLAP [5] es comunicar información de manera visual. Buendía Avalos [6], indica que las gráficas son excelentes formas de comunicación para plantear y resolver problemas, ejemplificar, ilustrar y comunicar resultados. Las encontramos en medios tan diversos como los periódicos, los libros de texto o los artículos de investigación. Por eso la educación tiene entre sus objetivos enseñarnos a hacer e interpretar gráficas.

Este software [7] también, posibilita la resolución de tareas de evaluación continua donde el alumno puede practicar los contenidos vistos en teoría, consiguiendo con su uso la autocorrección de ejercicios.

Nuestra práctica docente permite señalar que siempre al utilizar las herramientas tecnológicas en los procesos de enseñanza – aprendizaje, debe estar acompañada del background teórico que posibilita: inferir si resultados obtenidos mediante la tecnología son correctos, y generar fundamentos para nuevos productos informáticos. Cuando los actuales no responden a las necesidades presentes. En esta línea se sustenta la propuesta de buscar herramientas teóricas que luego se puedan implementar con uso del software para abordar su imposibilidad de realizar algunas tareas. En concreto nos focalizaremos en la derivación de expresiones regulares.

Para ello se propone adicionar al software desarrollos teórico-prácticos que permitan desde las reglas de la gramática del Autómata Finito Determinístico (AFD) y su Expresión Regular (ER), obtener las reglas de la gramática del AFD que corresponde a la derivada de la ER.

Lo cual permitiría el trabajo de la derivada de la ER con el software en cuestión. Logrando explorar con herramientas de visualización de modelos, nuevos ambientes de enseñanza-aprendizaje, que no se reducen a los meramente simbólico- algebraicos

El uso de TIC, en este caso el software JFLAP, en el ámbito de enseñanza universitaria aporta en la gestión educativa, espacios para apropiación de contenidos y de autoevaluación del alumno. Su colaboración en la ejecución de la asignatura permite obtener nuevas metodológicas de trabajo [8] que no limiten las prácticas áulicas sólo al uso de recursos tradicionales.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esencialmente la línea de investigación principal y específica en que se enmarca este trabajo, es el abordaje de la lingüística de los AFD de ER y sus derivadas (tratados con JFLAP ó con herramientas teórico-prácticas), que permitan realizar algoritmos, para implementar el software cuando el desarrollo de estos resulte complejo de trabajar en forma analítica ó no pueda realizar alguna tarea.

Entre los objetivos se plantean:

- Reconocer y potenciar las capacidades del software para dar respuesta a la representación gráfica del AFD de la derivada de una ER.
- Desarrollar algoritmos que generen la gramática del AFD de la derivada de una ER.
- Ensamblar los algoritmos desarrollados con la capacidad de visualización gráfica y producción simbólica-algebraica del software.

Para lo cual se han planificado las siguientes etapas:

a) Revisión de herramientas que se pueden obtener con el software desde el grafo de un AFD.
b) Investigación bibliográfica para determinar si se puede dar desde la derivada de una ER la gramática del AFD asociado, vinculando con la gramática del AFD de la ER objeto de trabajo.
c) Estudio de cómo determinar desde la gramática de un AFD y su ER asociada, las reglas de producción de

la gramática del AFD de la derivada de la ER.

d) Elaboración de algoritmo/s que vinculen ambas gramáticas.
e) Integración del/los algoritmo/s con el software JFLAP.

Los integrantes de este proyecto participan, de las cátedras de: Matemática Discreta, Teoría de Autómatas y Complejidad y de Algoritmos Numéricos. Por lo que, las áreas de interés se mezclan adecuadamente con este proyecto de investigación sus inquietudes y el trabajo académico habitual.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto se encuentra en sus primeras etapas y, hasta el momento, se han conseguido realizar las tareas:

- Se ha identificado que el software JFLAP no realiza la tarea de representar desde el grafo del AFD asociado a una ER, el grafo del AFD ligado a la ER obtenida de una derivada.
- Se ha comenzado el estudio bibliográfico para ver si existe solución ya realizada a lo planteado.
- Se han iniciado desarrollo teóricos, en caso de que no exista bibliografía desarrollada sobre la temática.
- Como resultado de proyectos llevados a cabo en temas contiguos a esta línea de trabajo, se han realizado varias presentaciones y publicaciones en congresos nacionales del área de investigación en Computación y en el área de Educación Matemática [9], [10].

Se espera contribuir en el área de la programación obteniendo un algoritmo útil y ejecutable para anexar a la tarea que se hace con JFLAP, logrando el ensamble final con el software. Una vez alcanzado el objetivo se revisará que el mismo sea funcional y de baja complejidad computacional.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigación está formado por cuatro docentes- investigadores, participa del proyecto el ayudante alumno de la asignatura, con lo cual su intervención le ayuda a aprender a realizar actividades de investigación, y cómo integrarse en un equipo de investigación de cátedra, existente.

Además, la producción científica realizada se puede volcar a temas de actualización a la asignatura, con lo cual los alumnos que cursan la materia también se retroalimentan de las investigaciones de cátedra.

5. REFERENCIAS

[1] A. Pelliccia, M. Vasta, *Aprender a Aprender, Ciencia e investigación*, Tomo 66 Nº 5, 2016

[2] L. Lluch Molins, Ma. Portillo Vidiella, *La competencia de aprender a aprender en el marco de la educación superior*, Revista Iberoamericana de Vol. 78, No. 2, pp. 59-76, ISSN: 1022-6508 / ISSNe: 1681-5653 Organización de Estados Iberoamericanos (OEI/CAEU) ,2018

[3] Ph. Meirieu, *Aprender, sí. Pero ¿Cómo?*, Octaedro, Madrid ,1992.

[4] M. Maggio, *Enriquecer la enseñanza. Los ambientes con alta disposición tecnológica como oportunidad.* Editorial Paidós, Buenos Aires, 2012.

[5] http://www.jflap.org/ consultado el 21/05/2018

[6] G. Buendía Avalos, *¿Como ves?* Revista de divulgación de la Ciencia de la UNAM.México. Año 17. No. 203. Octubre 2015.

[7] S.Rodger, T.Finley, *JFLAP: An Interactive Formal Languages and Automata Package*. Jones and Bartlett Publishers, Inc; Edición: Computer Scienc, 2006.

[8] A. Zangara, *Uso de nuevas tecnologías en la educación: una oportunidad para fortalecer la práctica docente*. Revista Puertas Abiertas - ISSN 1853-614X.

[9] E. Oliva, N. Alonso, *Derivadas de expresiones regulares-lenguajes asociados* Actas de XIII Congreso Argentino de Educación Matemática. Argentina. 2018.

[10] E. Oliva, M. Diaz Ogás, A.L. Molina,N. Alonso, *Lenguajes regulares y autómatas finitos determinísticos asociados. Discusión de implementación con Jflap.* Anales del 2do Congreso Internacional de Ciencias de la Computación y Sistemas de Información. Argentina. 2018 [2] P. Pacheco, *An introduction to parallel programming*. USA: Elsevier, 2011.

.

ACCESIBILIDAD EN ENTORNOS VIRTUALES DE ENSEÑANZA Y APRENDIZAJE

Adó Mariana[1], Arnaiz Lucas[2], Lencina Paula[1], Maccarino Christian[2], Ochipinti Pedro[1], Rodríguez Marina[1], Sarobe Mónica[1], Serrano Eliana[1]

Instituto de Investigación y Transferencia en Tecnología (ITT)[3]
Comisión de Investigaciones Científicas (CIC)
Escuela de Tecnología (ET)
Universidad Nacional del Noroeste de la Provincia de Buenos Aires (UNNOBA)

Sarmiento N° 1119 3er Piso, Junín (B) – TE: (0236) 4477050 INT 11610

{mariana.ado, paula.lencina, christian.maccarino, marina.rodriguez, eliana.serrano, pedro.ochipinti, monica.sarobe, lucas.arnaiz}@itt.unnoba.edu.ar

RESUMEN

La integración de la Tecnología de la Información y el Conocimiento (TIC) en las universidades, así como también la implementación de los Entornos Virtuales de Enseñanza y Aprendizaje (EVEA), abrieron una gran variedad de posibilidades de aprendizaje y acceso a la información. En consecuencia, gran cantidad de personas que antes no accedían a las instituciones de educación superior por razones de variada índole, como, por ejemplo: culturales, socio-económicas, políticas, geográficas, entre otras; hoy tienen la posibilidad de ingreso a las mismas.

Pero, del mismo modo, al aumentar las posibilidades para muchas personas, se aumentó la brecha digital entre aquellas personas que tienen mayores posibilidades en el acceso a la educación, a la información y a la tecnología, y la "gran minoría de personas para quienes la desigualdad de oportunidades en el acceso a la información y al conocimiento no les permite participar en igualdad de condiciones en la sociedad de la información" [1].

Esta línea de investigación, aborda el problema de la brecha digital existente para las personas ciegas, sordas y con visión o audición reducida, debido a la incorporación de los EVEA en las instituciones de educación superior y cómo éstas interactúan en ese contexto.

En particular, se establece como objetivo el desarrollo de una metodología de trabajo que cumpla con la legislación, los requisitos y estándares de accesibilidad vigentes y aprobados en nuestro país, para implementarse en el EVEA de la universidad [12] [13]. De esta manera, se espera disminuir la brecha digital existente entre aquellas personas que no poseen igualdad de acceso a la tecnología y la información, y las que sí; introduciendo una perspectiva de equidad e igualdad, a la hora de implementar nuevas políticas universitarias. [8][9]

---

[1] Docente Investigador - ITT
[2] Becario PROMINF - ITT
[3] ITT - Centro Asociado CIC

Palabras clave: Accesibilidad, Accesibilidad Web, Discapacidad, Entornos Virtuales de E-A, Inclusión, Brecha Digital

CONTEXTO

En el marco del proyecto de investigación "Informática y tecnologías emergentes" que se desarrolla en la UNNOBA, con lugar de trabajo en el ITT, se continúa trabajando con la línea de investigación "Accesibilidad en EVEA y su impacto en la brecha digital".

1. INTRODUCCIÓN

Son muchas las discapacidades que pueden limitar el uso de las TIC. Como principal solución a estos problemas existen herramientas, tanto de hardware como de software, que sustituyen la funcionalidad que no tienen los usuarios con discapacidad. Para garantizar el acceso a cualquier recurso tecnológico a todas las personas, incluidas las que sufren algún tipo de discapacidad, no es suficiente con desarrollar tecnologías asistidas adecuadas para cada tipo de discapacidad. También es necesario que los propios recursos tecnológicos estén diseñados para que puedan ser accedidos a través de estas ayudas técnicas [10].

En la última década, la educación superior ha puesto como una de sus preocupaciones principales el derecho a una educación de calidad para todas las personas, partiendo de los principios universales de diversidad, igualdad y equidad [11].

En este contexto, Mareño Sempertegui y Torrez definen accesibilidad como:

"el requisito fundamental que deben cumplir los espacios, ámbitos, servicios, bienes, así como los objetos o instrumentos, herramientas y dispositivos, para poder ser utilizables por todas las personas de manera: autónoma, es decir, en forma independiente, no subordinada al auxilio de otra; segura, o sea, libre de todo peligro o riesgo; confortable, es decir, de manera cómoda sin inconvenientes, obstáculos o restricciones en su uso, y con un mínimo de esfuerzo; y equitativa, es decir, con equidad entendida como la cualidad que consiste en atribuir a cada persona aquello que le corresponde por derecho" [2].

Además, se define accesibilidad web como "la posibilidad de que la información de la página Web, puede ser comprendida y consultada por personas con discapacidad y por usuarios que posean diversas configuraciones en su equipamiento o en sus programas" [3]. Es decir, Mareño Sempertegui et al. [2] afirman que la accesibilidad web: "involucra aquellos aspectos relacionados con la codificación y la presentación de información en el diseño de un sitio web, que van a permitir que cualquier persona, sin importar sus características biofísicas, pueda percibir, entender, navegar e interactuar de forma efectiva con la web, así como crear y aportar contenidos, independientemente del tipo de hardware, software, infraestructura de red que utilice, idioma, condición socio-económica, cultura o localización geográfica" [2].

Según Pugliese [4], los informes de la Comisión Universitaria sobre Discapacidad y Derechos Humanos de Argentina, espacio multiactoral conformado por gran parte de las universidades nacionales del país, sostienen que uno de los mayores obstáculos para la equiparación de oportunidades en el ámbito universitario, en lo que refiere específicamente a los entornos virtuales, es que la mayoría de los sitios web de las universidades presentan barreras de accesibilidad, lo que hace difícil e incluso imposible su utilización.

Algunos de los problemas de accesibilidad web a los que se enfrentan las personas ciegas, sordas o con problemas de visión o audición al navegar por las páginas o aplicaciones web, son que las imágenes y los enlaces gráficos no poseen descripciones en texto alternativo, o si lo poseen, no están definidos de forma adecuada; por lo que no pueden ser interpretados por lectores de pantalla de manera correcta, ni pronunciados por equipos con reconocimiento de voz. A eso se suma que los videos no están subtitulados. Según González Campo [5], el texto está codificado de forma rígida, donde los usuarios están imposibilitados para seleccionar su propio tamaño o estilo de letra, para hacerla más legible; la multimedia y otras representaciones interactivas, no poseen ninguna alternativa no gráfica para entender su contenido.

Hilera González [6], define un EVEA accesible como aquel producto y servicio de e-learning que es comprensible, utilizable y practicable por los estudiantes que posean los conocimientos previos establecidos, independientemente de que tengan algún tipo de discapacidad.

Mareño Sempertegui [7], plantea que se habla de personas en situación de discapacidad para referirse a aquellos sectores poblacionales tradicionalmente nominados como lisiados, impedidos, minusválidos, deficientes, discapacitados o personas con discapacidad, según diversos momentos históricos y perspectivas teóricas de análisis. Se concibe a la discapacidad como una situación construida social y culturalmente, una creación humana que todos los actores sociales producimos día a día a partir de actitudes, decisiones, acciones y omisiones fundadas en la naturalización de criterios de normalidad. Es la situación construida, que involucra aspectos físicos, sociales, culturales, políticos, económicos y actitudinales la que limita o no limita, la que permite o no permite, la que posibilita o no posibilita; independientemente de las características personales o corporales de las personas que participen de ella.

La educación superior actual se encuentra en constante evolución en la sociedad del conocimiento. La adopción de TIC en las actividades académicas se refleja en la oferta de cursos de modalidad virtual y también en el uso de entornos virtuales en los planes de modalidad presencial. Sin embargo, los requisitos de accesibilidad de las plataformas y los cursos virtuales no se han tomado en cuenta en el proceso formativo.

En este marco, se considera que la ausencia de condiciones de accesibilidad en cualquier entorno virtual constituye una forma de discriminación, al no garantizar la igualdad real de oportunidades y trato, en este caso, una exclusión digital, educativa y social [19]. Mareño Sempertegui y Torrez [2], exponen que la inexistencia de condiciones de accesibilidad web, genera barreras en la comunicación y en el acceso a la información y al conocimiento, que constituyen barreras institucionales no explícitas e invisibles.

En nuestro país, se ha comenzado a tener en cuenta la accesibilidad web en los últimos años, resultando en la definición de un marco legal y estándares aprobados, lo cual representa un magnífico beneficio para estudiantes con discapacidad [18].

En este sentido, es necesario que en las universidades se reconozca "la importancia de incorporar condiciones de accesibilidad en los entornos virtuales de las Instituciones de Educación Superior (IES), como condición

necesaria para favorecer la expansión de la equidad y el acceso a la educación superior y, por ende, fortalecer los procesos institucionales de democratización del conocimiento y la información" [3].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Esta línea de investigación aborda el problema de la brecha digital existente para las personas con discapacidad visual o auditiva, debido a la incorporación de EVEA en las instituciones de educación superior. En particular, se establece como objetivo desarrollar una metodología de trabajo que cumpla con la legislación, los requisitos y estándares de accesibilidad vigentes y aprobados en nuestro país, para luego implementarla en el EVEA de la UNNOBA [12] [13].

Como objetivos específicos se definen:

1. Analizar y hacer un diagnóstico de la accesibilidad del EVEA de la UNNOBA.
2. Definir una metodología que establezca un modelo de trabajo para el cumplimiento de requisitos y estándares de accesibilidad en el contexto de la formación virtual.
3. Diseñar el EVEA en concordancia con la metodología propuesta.
4. Implantar el EVEA accesible en la UNNOBA.

Al tratarse de un trabajo de investigación tecnológica [16], el producto final será un software, en este caso se trata de un EVEA accesible. Particularmente, se adoptará la metodología de Investigación Basada en Diseño (IBD) [14] [15] que posee dos etapas principales: la primera es investigar para crear un nuevo producto, realizar el primer prototipo y sus sucesivas mejoras; y la segunda aportar conocimiento en forma de principios que contribuyen a nuevos procesos de diseño.

En el marco de la primera etapa de la metodología IBD se utilizará el Pensamiento de Diseño [17], el cual está centrado en el usuario y en la generación de ideas innovadoras para entender y dar solución a las necesidades reales de los clientes o usuarios finales, utilizando un enfoque multidisciplinario.

3. RESULTADOS OBTENIDOS / ESPERADOS

Se espera, desarrollar una metodología de trabajo acorde a la legislación vigente de accesibilidad en nuestro país, para el diseño y desarrollo de los EVEA y del material educativo para los mismos, que nos permita confeccionar políticas universitarias de inclusión e igualdad de acceso para todas las personas, independientemente de sus condiciones físicas, culturales, socio-económicas, políticas, geográficas, entre otras.

En cuanto al cumplimiento de la segunda etapa de IBD, se espera poder aportar conocimiento al definir una metodología de trabajo.

Asimismo, se está programando una capacitación para el plantel docente y no docente de la universidad; y se aspira a la difusión y generación de conocimiento 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 tres investigadoras formadas del área de Sistemas y Tecnología aplicada a la educación, un investigador del área de Sistemas, dos investigadoras en formación y dos becarios de grado.

Asimismo, se espera que, durante el transcurso del año, se concrete una tesis de la "Maestría en Educación en Entornos Virtuales" de la Universidad Nacional de la Patagonia Austral.

5. BIBLIOGRAFÍA

[1] Sáenz Espitia, J.G. (2008). *Accesibilidad en las aulas de la web 2.0.* Bdigital Portal de Revistas UNAL. Volumen 1. Recuperado de: https://revistas.unal.edu.co/index.php/email/article/view/13970/14737

[2] Mareño Sempertegui, M., Torrez, V. (2013). *Accesibilidad en los entornos virtuales de las instituciones de educación superior universitarias.* Virtualidad, Educación y Ciencia.Vol. 4, Núm. 7. Recuperado de: https://revistas.unc.edu.ar/index.php/vesc/article/view/6176

[3] Ley 26.653, Accesibilidad de la Información en las Páginas Web, 26 de noviembre de 2010. Recuperado de: http://servicios.infoleg.gob.ar/infolegInternet/anexos/175000-179999/175694/norma.htm

[4] Pugliese, J. C. (2005). *La Integración de las Personas con Discapacidad en la Educación Superior en la República Argentina.* Buenos Aires, Argentina: Secretaría de Políticas Universitarias, Ministerio de Educación, Ciencia y Tecnología.

[5] González Campo, C. H. (2009). *Herramientas Web 2.0 y accesibilidad a sitios Web para la apropiación social de conocimiento en una ciudad educadora.* Revista Universidad EAFIT. Vol. 45. No. 153. pp. 26-37. Recuperado de: http://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/76/75

[6] Hilera González, J.R. (2012). *El Proyecto ESVI-AL.* Recuperado de: http://www.esvial.org/wp-content/files/boletin1_A4.pdf

[7] Mareño Sempertegui, M. (2010). *Inclusión Educativa en la Universidad Nacional de Córdoba: el desafío de aplicar los principios del Diseño Universal en la gestión de políticas universitarias.* Ponencia presentada en: Primeras Jornadas Nacionales de Accesibilidad y Diseño Universal. Córdoba, Argentina: Universidad Nacional de Córdoba.

[8] Cambours de Donini, A., Gorostiaga J. (2016). *Hacia una universidad inclusiva. Nuevos escenarios y miradas.* Aique Grupo Editor.

[9] Sitio Web Oficial de la organización ESVI-AL (Educación Superior Virtual Inclusiva de América Latina), Recuperado de: http://www.esvial.org/

[10] Guenaga M., Barbier A. Eguíluz A. (2007). *La accesibilidad y las tecnologías en la información y la comunicación.* Trans Revista de Traductología No. 11. pp. 155-169. Recuperado de: http://www.revistas.uma.es/index.php/trans/article/view/3104/2867

[11] Cicerchia L., Addante P., Ahmad T., Lencina P., Russo C. (2014). *Accesibilidad Web con herramientas de Software Libre en UNNOBA Virtual.* Recuperado de: https://digital.cic.gba.gob.ar/bitstream/handle/11746/7087/Poster_Accesibilidad_Web_UNNOBA_CICERCHIA.pdf?sequence=1

[12] Entorno Virtual de Enseñanza y Aprendizaje (EVEA) de la UNNOBA, http://virtual.unnoba.edu.ar

[13] Sitio web oficial del entorno virtual de aprendizaje Moodle, https://moodle.org/

[14] Amiel, T., & Reeves, T. C. (2008). *Design-Based Research and Educational Technology: Rethinking Technology and the Research Agenda.* Educational Technology & Society, Vol. 11 No. 4, pp. 29–40.

[15] De Benito, B. y Salinas, J.M. (2016). *La investigación basada en diseño en Tecnología Educativa. RIITE.* Revista Interuniversitaria de Investigación en Tecnología Educativa, No. 0, 44-59

[16] Muñoz Razo, C. (2015). *Cómo elaborar y asesorar una investigación de tesis.* México: Pearson Educación.

[17] Design Thinking en español (2018). Fases del proceso de Design Thinking. Recuperado de: http://www.educacionyculturaaz.com/noticias/design-thinking-y-sus-tecnicas

[18] Sitio Web Oficial del Gobierno de la Nación Argentina. *Accesibilidad Web. Reglamentación.* Recuperado de: https://www.argentina.gob.ar/onti/accesibilidad-web

[19] Hilera González, J.R. (2012). *El Proyecto ESVI-AL.* Recuperado de: http://www.esvial.org/wp-content/files/boletin1_A4.pdf

WICC San Juan 2019
21° Edición del Workshop de Investigadores en Ciencias de la Computación

San Juan, 26 y 27 de Abril de 2019

Alfabetización con formatos múltiples de aprendizaje: Las TAC y los nuevos escenarios tecno-pedagógicos en la universidad

González, Liliana M[1]; Ureta, Laura; Marcovecchio, María José; Margarit, Viviana; Pontoriero, Francisco; Rossetti, Gabriela; Villodre, Silvia

[1] e-mail: lilianamirna@gmail.com

Grupo de investigación perteneciente al Gabinete de Tecnología e Innovación Educativa.
Dpto. de Informática - Dpto. de Geofísica y Astronomía. - Dpto. de Biología
Facultad de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de San Juan

Resumen

Las TIC forman parte de nuestra vida diaria y esta realidad es aún más notable entre los jóvenes, quienes mantienen lazos estrechos con ellas y no podrían pensarse sin ellas tanto en su vida social como académica. Las nuevas tecnologías atraviesan sus modos de pensar, aprender y conocer; he aquí la importancia de que las instituciones educativas, docentes y otros responsables de niños y jóvenes reconozcamos las tendencias culturales. Cabe preguntarnos, ¿Cuáles son las tendencias culturales respecto al uso de las Tecnologías que hacen nuestros jóvenes? ¿Existe interés en los docentes por reconfigurar las prácticas educativas y las aulas utilizando TAC? ¿El diseño de escenarios tecno pedagógicos con TAC favorece la construcción de conocimiento? El objetivo de nuestra investigación es identificar cuáles son los modos de uso de las TIC y TAC en carreras universitarias con el fin de generar conocimientos que aporten a la problemática de cómo repensar y reconfigurar las prácticas educativas.

El modelo TPACK constituye un marco teórico interesante para una integración eficaz de las tecnologías en la enseñanza. Metodológicamente, se optó por abordar el objeto de estudio con metodología cualitativa, específicamente un estudio descriptivo, lo que permitirá conocer los modos de uso de las TIC y las TAC en las prácticas educativas por los docentes en carreras de la FCEFyN.

Todos los docentes tenemos un desafío importante, es necesario que generemos nuevos escenarios tecno-pedagógicos que alberguen estrategias y propuestas educativas abiertas y flexibles con tecnologías que aporten potencialidades para Saber, Saber hacer y Saber ser.

Palabras clave: ALFABETIZACIÓN-TAC-ESCENARIOS TECNOPEDAGÓGICOS

Contexto

El Proyecto mencionado -aprobado por CICITCA- tiene como objetivo principal generar conocimientos que aporten a la problemática de ***cómo repensar y reconfigurar prácticas educativas con TAC en carreras universitarias.*** Se encuentra enmarcado en el "Gabinete de Tecnología e Innovación Educativa creado por Resolución 03/18 de la Facultad de Ciencias Exactas, Físicas y Naturales. El Gabinete se constituye en un espacio académico y de investigación cuyo objetivo es generar estrategias, metodologías, medios y materiales así como instrumentos tecnológicos, que contribuyan a la inserción de innovaciones tecnológicas en pos de la mejora de los procesos educativos de la Facultad de Ciencias Exactas, Físicas y Naturales de la Universidad Nacional de San Juan

1. Introducción

Las nuevas tecnologías atraviesan los modos de pensar, aprender y conocer de nuestros estudiantes; he aquí la importancia de que las instituciones educativas, docentes y todos los responsables de la formación de niños y jóvenes reconozcan las tendencias culturales. No podemos negar que nuestros jóvenes tienen el mundo en las yemas de sus dedos, y que la inmediatez e instantaneidad son factores muy presentes en sus actividades. Aquí es pertinente considerar la propuesta de Maggio (2012), quién expresa "[…] lo primero que debemos hacer es reconocer lo que los alumnos hacen" (p.153), preguntarnos qué son capaces de hacer, cuáles son sus preferencias tecnológicas, en qué actividades las usan y cómo usan las tecnologías, qué valor les asignan en sus procesos de estudio y producción. Esto significa reconocer el valor de los "[…] usos culturales y formas cognitivas del mismo modo como deberíamos tener en cuenta los estilos cognitivos propios de los sujetos culturales que son nuestros alumnos" (p.153) La sociedad de la información pone énfasis en la dimensión intersubjetiva de los aprendizajes en contraposición a las concepciones de aprendizaje propias de la sociedad industrial que centró su atención en la dimensión intrapsicológica del alumno. Aún coexisten diferentes concepciones y estructuras institucionales que siguen dando lugar a rupturas que problematizan la realidad que nos toca vivir. Área Moreira (2009) y Cabero Almenara (2011) coinciden en reconocer y plantear una problemática que se evidencia al momento de implementar las nuevas tecnologías en las aulas. Este último autor expresa "[…] encontramos una escuela articulada alrededor de diferentes siglos: escuelas con estructuras organizativas del siglo XIX, profesores del siglo XX y alumnos del siglo XXI". Hoy se da la naturalización de ciertas prácticas sociales y culturales que hasta hace poco tiempo no habían impactado en los vínculos sociales y en las formas de comunicación. Los jóvenes se ven beneficiados por el acceso a dispositivos personales móviles, lo que permite que estén "conectados" de forma continua. La participación de los estudiantes en redes sociales como Facebook, grupos de Whatsapp, foros de temas de su interés y otras formas de participación hace que se evidencie una actividad intensa con diferentes tecnologías.

El nuevo contexto socio-tecnológico requiere que el estudiante desarrolle capacidades que le permitan no sólo acceder a la información y al conocimiento mediante el uso de las tecnologías sino también adquirir, participar, comunicar y colaborar con el conocimiento en la red. Las nuevas formas de construcción de conocimiento en entornos digitales requiere de la integración de habilidades a fin de que el estudiante pueda buscar, encontrar, analizar, evaluar y utilizar eficazmente la información que necesita pero además transformar, comunicar y colaborar con sus pares en la resolución de problemas. A partir de ésto resulta evidente la necesidad de empoderar a nuestros alumnos con capacidades que los ayuden a mejorar su autonomía, la toma de decisiones, la identificación de sus propias necesidades.

El uso de estas tecnologías en la educación "por sí" no basta para un óptimo proceso de aprendizaje. Precisamente, por ello, nació hace unos años el término TAC (Tecnologías del Aprendizaje y del Conocimiento) haciendo referencia al uso de las TIC como herramienta formativa que inciden en la metodología y en la utilización de la tecnología dentro de las planificaciones educativas. En otras palabras, el término TAC destaca las nuevas posibilidades que las tecnologías abren a la educación, cuando éstas dejan de usarse como un elemento meramente instrumental. Su nueva función posibilita que "[…] el contexto socio-tecnológico genere un nuevo modelo de escuela que responda a las necesidades formativas de los ciudadanos". (Castañeda, Adell, 2013). De esta manera, las TAC se convierten en una "inclusión" de las tecnologías en la educación.

Teniendo en cuentas estas conceptualizaciones, y tomando como punto de partida la caracterización de las propuestas educativas con TIC, nos proponemos diseñar escenarios tecno-pedagógicos con TAC que favorezcan el desarrollo de competencias en nuestros alumnos, como sujetos culturales en un paradigma socio-tecnológico. Tales habilidades apuntan a la formación del rol de auto-gestor del alumno en su aprendizaje.

2. Líneas de Investigación, Desarrollo e Innovación

- Impactos de la tecnología en los modos de conocer en el marco de escenarios tecno-pedagógicos con TAC en carreras universitarias.

3. Resultados Obtenidos/Esperados

- Diseño e Implementación de un escenario tecno-pedagógico para el desarrollo de la asignatura Inglés nivel III de la carrera Licenciatura en Biología. F.C.E.F.N.

- Experiencia piloto: grupo experimental con alumnos del segundo año de las carreras de Profesorado y Licenciatura en Sociología de la F.C.S. utilizando mensajería instantánea, Whatsapp.

- Exposiciones y Publicaciones

- González, Liliana; Ureta, Laura y Balmaceda Silvina. "Espacios dialógicos para la construcción de conocimientos conocimientos en procesos de formación de postgrado. Virtualización de la Educación Superior. III Jornadas Nacionales y V Jornadas de la UNC. Experiencias e Investigación en Educación a Distancia y Tecnología Educativa. 2018. Universidad Nacional de Córdoba.

- Ureta, Laura. "Prácticas pedagógicas con TIC para la escritura de un *abstract* en lengua extranjera". II Congreso Internacional de Lenguas-Migraciones-Culturas. Facultad de Lenguas, Universidad Nacional de Córdoba, mayo 2018.

- Llarena, Myriam y Villodre Silvia. Los entornos Personales en los Espacios Virtuales de Aprendizaje.2018. Virtualización de la Educación Superior - III Jornadas Nacionales y V Jornadas de la UNC. Experiencias e Investigación en Educación a Distancia y Tecnología Educativa. Universidad Nacional de Córdoba

- Ureta, Laura; Díaz Ogas, Fernando; Becerra, Sofía, Villalobos Perna, Priscila. "Construcción de conocimientos de las Ciencias Biológicas en Inglés con tecnologías". IV Jornadas de Investigación Educativa y III Jornadas de Prácticas de la Enseñanza del Profesorado en Ciencias Biológicas. Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, noviembre 2018. ISBN: en trámite.

- Marcovecchio, María José y Margarit, Viviana. "M-LEARNING: construyendo conocimiento asistido por las nuevas tecnologías para la enseñanza del inglés en el ámbito universitario". XX Congreso Internacional EDUTEC y El Centro de Investigación e Innovación en Educación y

TIC (CIIET) de la Universidad de Santiago de Chile. Comunicación publicada en el Libro de Resúmenes: "Educación y Tecnología: una mirada desde la Innovación e Investigación". Enero de 2018. ISBN: 978-956-393-378-9.

- Marcovecchio, María José y Margarit, Viviana. Exposición de la ponencia: *M-learning: el Whatsapp como herramienta TIC aplicada en aprendizaje del inglés* durante el XXIV Encuentro Plurilingüístico "Más lenguas…más Empatía", organizado por la Universidad Católica de Cuyo, a través de su Dpto. de Idiomas dependiente de la Fac. de Educación, el Colegio Central Universitario "Dr. Mariano Moreno" (UNSJ), a través de su Dpto. de Lenguas Extranjeras, y la Fundación Instituto Alemán - Goethe Zentrum San Juan. Agosto de 2018.

4. Formación de Recursos Humanos

El equipo de trabajo está conformado de la siguiente manera:

- Directora: González, Liliana Mirna
- Co-directora: Ureta Laura
- Investigadores:
  - Pontoriero, Francisco
  - Margarit Viviana
  - Marcovecchio, María José
  - Rossetti Beiram, Gabriela Maribel
  - Villodre, Silvia Lydia

Se propone volcar los resultados en actividades de Postgrado, tales como:

- Diseño e implementación de la Diplomatura "Educación y nuevas tecnologías en tiempos de convergencia tecnológica". Aprobada por Ordenanza Nº 5 CD-FCEFN, Prof. Responsable: Mgter Liliana Mirna González. 1º cohorte mayo -- Diciembre 2015. 2º cohorte mayo -- Diciembre 2016. 3º cohorte mayo -- diciembre 2017. 4° cohorte agosto 2018
- Co-dirección/Asesoramiento de Tesis de Maestría en Informática. F.C.E.FyN de la UNSJ.
- Docente de la Maestría en Informática. F.C.E.FyN de la UNSJ. Materia: Educación y Tecnologías.

5. Bibliografía

- Canales Cruz Alejandro (2014) Hacia un nuevo diseño para el aprendizaje: escenarios educativos para la Web 2.0. Revista de Innovación Educativa. Vol. 6, Núm. 2

- Casablancas S. (2012) Tecnologías digitales en las aulas: roles docentes, de los estudiantes y del conocimiento" para las IV Jornadas de TIC e Innovación en el Aula - UNLP https://jornadas.ead.unlp.edu.ar/ consultado 1/12/17.

- Enríquez, Silvia. (2012) "¿TIC o TAC? ¿Cómo debe ser la alfabetización digital de los docentes?" Ponencia presentada en el congreso virtual Eduq@2012, junio de 2012. Pendiente de publicación en actas. Traducción: Emilio Quintana. Grupo Nodos Ele: www.nodosele.com "Ubiquitous Learning. Exploring the anywhere/anytime possibilities for learning in the age of digital media". Editorial Universidad de Illinois, 2009

- Litwin E. (2005) Tecnologías educativas en tiempos de Internet. Amorrortu. Argentina.

- Lozano, R. (2011) "Las 'TIC/TAC': de las tecnologías de la información y comunicación a las tecnologías del aprendizaje y del conocimiento". En http://www.thinkepi.net/las-tic-tac-de-las-tecnologias-de-la-informacion-ycomunicacion-a-las-tecnologias-del-aprendizaje-y-del-conocimiento.Último ingreso mayo de 2012.

- Maggio M. (2012) Enriquecer la Enseñanza. Paidós. Argentina

- Castañeda, L. y Adell, J. (2013). La anatomía de los PLEs. En L. Castañeda y J. Adell (Eds.), Entornos Personales de Aprendizaje: Claves para el ecosistema educativo en red (pp. 11-27). Alcoy: Marfil

- Koeheler, Matthew y Punya Mishra (2006), “Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge”, Teachers College Record, 108(6), 1017-1054. Disponible en inglés en: http://punya.educ.msu.edu/publications/journal_articles/mishrakoehler-tcr2006.pdf (última consulta: septiembre de 2014).
- Lorenzo-Lledó, Alejandro (2018) Innovación en el aprendizaje desde el diseño tecno-pedagógico

- Sevilla H., Tarasow F. y Luna M. (coord.), (2017) Educar en la era digital. Docencia, tecnología y aprendizaje. Pandora Impresores, Jalisco, México.Carlino P. (2005) Escribir, leer y aprender en la universidad: una introducción a la alfabetización académica. F.C.E. Argentina

Análisis de perfiles de rendimiento académico mediante técnicas de minería de datos y análisis de datos multivariados

Maria Paula DIESER[(1)], María Cristina MARTÍN[(1)(2)], Lorena Veronica CAVERO[(1)], Sofía FUNKNER[(1)], Laura WAGNER[(1)]

[(1)] Facultad de Ciencias Exactas y Naturales, Universidad Nacional de La Pampa
[(2)] Departamento de Matemática, Universidad Nacional del Sur

{pauladieser, maritamartin}@exactas.unlpam.edu.ar

RESUMEN

La deserción estudiantil, especialmente en los primeros años de la carrera, es una preocupación presente y constante en todas las Instituciones de Nivel Superior. Para un tratamiento efectivo y eficaz del problema, resultan indispensables la detección temprana de estudiantes en situación de riesgo en términos de abandono o retraso en el alcance del grado, y el diseño e implementación de un plan de acción consecuente.

La Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa no es ajena a esta realidad. En consecuencia, la línea de investigación aquí presentada, propone estudiar y aplicar distintos métodos que ofrece la Minería de Datos y el Análisis de Datos Multivariados 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 el abandono de los estudios.

Palabras clave: minería de datos educativos, análisis de datos multivariados, rendimiento académico, deserción universitaria.

CONTEXTO

La línea de investigación que aquí se describe, se instala dentro de un Proyecto de Investigación más amplio, acreditado y financiado por la Facultad de Ciencias Exactas y Naturales (FCEyN) de la Universidad Nacional de La Pampa (UNLPam). Dicho Proyecto se deriva de tareas de investigación desarrolladas en la Institución durante el periodo 2014 – 2017, y vinculadas con el estudio y aplicación de métodos multivariados discriminantes y de clasificación, 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.

De las investigaciones realizadas, surge el campo de la educación como un terreno propicio para la aplicación de técnicas de Minería de Datos (MD), que pueden complementarse con otras propias del Análisis de Datos Multivariados (ADM). Además, tales métodos pueden nutrirse con 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 particular, la línea de investigación que aquí se presenta, iniciada en 2018, tiene por objetivo general estudiar y aplicar distintos métodos que ofrece el ADM y la MD sobre los datos registrados en SIU Guaraní 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 resultar en abandono de los estudios.

1. INTRODUCCIÓN

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.

Al mismo tiempo, 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 de información 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.

La MD, en combinación con el ADM, reúnen un conjunto de técnicas capaces de modelizar y resumir dicha información, facilitando su comprensión y ayudando a la toma de decisiones en situaciones futuras (Cabena et al. 1998; Hernández Orallo et al., 2004). En particular, la Minería de Datos Educativos (MDE) se presenta como un área iluminada por diferentes disciplinas, relativamente reciente y de crecimiento notable, que se ocupa del desarrollo, la investigación y la 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 (Romero & Ventura, 2010).

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 académico de los 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, 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).

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 enmarca en un Proyecto más amplio y tiene por objetivo general estudiar distintas técnicas de la MD y el ADM utilizadas en el campo educativo para la obtención de modelos de estudiantes que permitan identificar situaciones de riesgo de deserción o abandono.

Las técnicas estudiadas serán aplicadas sobre los datos registrados en el sistema de gestión de información estudiantil (SIU Guaraní) de la FCEyN (UNLPam) mediante rutinas desarrolladas en el lenguaje de programación R y, eventualmente, otros programas existentes para el procesamiento de datos (*e.g.* RapidMiner, STATGRAPHICS, STATA, SPSS, S-PLUS, NTSys, STATISTICA).

Así, los resultados obtenidos serán evaluados y comparados de manera que los mejores modelos sean utilizados 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

A lo largo del primer año se han desarrollado las tareas que se enumeran a continuación:

- Revisión sistemática de bibliografía referida a experiencias desarrolladas en el ámbito de la Educación Superior que utilicen la MDE y el ADM para identificar modelos que describen la trayectoria académica de estudiantes y patrones de deserción o abandono, poniendo especial atención a las técnicas y *software* utilizados, los atributos considerados, y aquéllos que se vinculan de manera significativa con el rendimiento académico.
- Estudio pormenorizado de las técnicas empleadas en las investigaciones empíricas revisadas, y otras de eventual utilidad, apoyado por el desarrollo de prácticas y aplicaciones sencillas mediante el lenguaje R y *software* estadístico de utilidad.
- Análisis de datos censales disponibles en SIU Guaraní (historia académica, atributos personales, y de índole social y económica) correspondientes a 9187 registros de estudiantes que han asentado su ingreso a la Institución a partir del año 2001 y hasta 2018.

A partir de estas tareas previas, y a fin de alcanzar los objetivos propuestos, se espera en los próximos años llevar adelante las siguientes acciones, aún pendientes:

- Aplicación de las técnicas estudiadas, y otras que pudieran emerger como potencialmente útiles, sobre los datos provenientes de SIU Guaraní de la FCEyN (UNLPam), previo desarrollo de técnicas de preprocesamiento (limpieza, selección de variables, y la transformación o combinación de éstas) que permitan obtener una vista minable de los datos recopilados.
- Evaluación y comparación de 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 total 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, trabajan tres docentes investigadoras cuya formación de base corresponde al campo de la matemática o la educación matemática.

- Dos de ellas han 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. Una de ellas proyecta trabajar en temas de autorregulación del aprendizaje y su impacto en el rendimiento académico de estudiantes universitarios. Los resultados derivados del proceso de investigación correspondiente podrían abrevar los objetivos propuestos para la línea de investigación descripta, y ampliar el rango de datos y atributos estudiados originalmente.
- La tercera ha comenzado sus estudios de Doctorado en Estadística en la Universidad Nacional de Rosario y proyecta realizar su trabajo de tesis doctoral en el área del ADM. Los conocimientos alcanzados podrían nutrir los procesos desarrollados al interior de este Proyecto y, en particular, los propios de la línea de investigación que aquí se presenta.

El equipo de trabajo cuenta también con la participación de una estudiante avanzada de Licenciatura en Matemática que ha orientado su formación específica en temas de estadística aplicada. Como tal, ha obtenido una Beca de Iniciación a la Investigación otorgada por la UNLPam, para desarrollar el proyecto "Los Árboles de Decisión aplicados al análisis de Datos Educativos" bajo la dirección de la Lic. Dieser. Se espera que los resultados alcanzados durante el periodo de Beca permitan a la postulante especializarse en una de las técnicas abordadas en esta línea de investigación y eventualmente, iniciar estudios de postgrado en estos temas u otros vinculados.

En cualquier caso, se espera originar y organizar un plantel humano (docentes, investigadores, y estudiantes) que aporte información de utilidad para generar políticas adecuadas en relación al ingreso y permanencia de estudiantes en la FCEyN (UNLPam). Asimismo, se considera la posibilidad de desarrollar otros temas de interés relacionados con las investigaciones que se realizan en el marco de esta línea y del Proyecto en general, que pudieran surgir como consecuencia de replanteos o de consultas atendidas en el asesoramiento estadístico a otros investigadores o instituciones del medio.

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.

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.

Hernández Orallo, J., Ramírez Quintana, M. J., & Ferri Ramírez, C. (2004). *Introducción a la Minería de Datos*. Madrid: Pearson Prentice Hall.

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.

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.

Romero, C. & Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Syst. Appl.*, 33(1):135–146.

Con formato: Inglés (Estados Unidos)

Con formato: Inglés (Estados Unidos)

Con formato: Inglés (Estados Unidos)

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.

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.

Aplicación móvil de test de orientación vocacional para la Universidad Nacional de Catamarca como herramienta para ayudar a disminuir la deserción universitaria y promocionar las carreras

Lazarte, Ivanna[1,2]; Doria, Vanesa[1]; Flores, Carola[1]; Moreno, Juan Pablo[1]; Fernández, Natalia[1]; Poliche, Valeria[1]

[1] Departamento Informática/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
[2] Departamento de Formación Básica/Facultad de Tecnología y Ciencias Aplicadas/Universidad Nacional de Catamarca
Maximio Victoria Nº 55 - C.P: 4700 - San Fernando del Valle de Catamarca.
Tel: +54 (03834) 435112
ilazarte@tecno.unca.edu.ar, vanesadoria@gmail.com, carolaflores@tecno.unca.edu.ar

Resumen

El presente artículo expone lo realizado en un proyecto de I+D, que tiene por objetivo proporcionar a la Dirección de Orientación Vocacional, dependiente de la Secretaria de Bienestar Universitario y Asuntos Estudiantiles de la Universidad Nacional de Catamarca (UNCA), una herramienta tecnológica que automatice el test de orientación vocacional y que sirva de instrumento para ayudar a los estudiantes del último año del Nivel Secundario a elegir una carrera, basada en información sobre sus gustos, intereses, habilidades y aptitudes. El propósito de contar con esta herramienta es, por un lado, promocionar las carreras que se dictan en las distintas unidades académicas de la UNCA y, por otro lado, ayudar a disminuir la deserción de los estudiantes de los diferentes programas académicos de la UNCA.

Palabras clave: Orientación vocacional, Promoción de carreras, Deserción universitaria, Aplicación móvil

Contexto

Este trabajo de investigación se lleva a cabo dentro del Laboratorio de Tecnologías de Información y las Comunicaciones (LaTIC's), en el ámbito de la Facultad de Tecnología de la UNCA y se encuentra enmarcado dentro del Proyecto de I+D "Desarrollo de una aplicación móvil de test de orientación vocacional para la Universidad Nacional de Catamarca". Bajo la dirección de la Dra. Ivanna Lazarte y la co-dirección de la Lic. Valeria Póliche, cuenta con aval institucional otorgado mediante Resolución Rectoral N° 494/2018. La línea de investigación está orientada al estudio del efecto que tiene la orientación vocacional en la elección de carrera, como así también la incidencia que tiene la elección de la carrera de acuerdo al resultado del test en el éxito escolar del alumno.

Introducción

La deserción universitaria es uno de los grandes problemas que afecta a muchas universidades de América Latina [1, 2, 3, 4]. Las causas de la deserción son múltiples y variadas por lo que esta problemática debe analizarse teniendo en

cuenta factores de índole personal, socioeconómico y académico, entre otros. Los factores académicos tienen que ver, entre otras causas, con la deficiente orientación vocacional, la insatisfacción con la calidad de la carrera, y la insuficiente preparación para ingresar a una carrera [4, 5, 7].

La orientación vocacional es reconocida como un instrumento de gran valor para la retención de estudiantes en los centros educativos, particularmente en la Universidad [1, 5, 6, 8, 9, 10].

La orientación vocacional es un proceso de ayuda para el estudiante, destinado a conseguir una comprensión adecuada de las distintas opciones profesionales que existen en el mundo de la educación, eligiendo aquella que cumpla con sus intereses y objetivos personales [9, 10]. Se considera que un proceso de orientación vocacional es satisfactorio cuando colabora para que el estudiante elija el camino a seguir lo más acertadamente posible, reduciendo frustraciones en las elecciones profesionales y vocacionales, y contribuyendo a alcanzar una mejor calidad de vida [11].

En el I Congreso Iberoamericano de Orientación, celebrado en Argentina en el año 2004, se propuso la participación activa que deben tener las Universidades para que las actividades de orientación vocacional sean definitivamente incorporadas a las actividades normales y propias de las mismas. En este sentido, la UNCA, a través de la Dirección de Orientación Vocacional, ofrece un proceso institucional de Orientación o Reorientación Vocacional, cuya modalidad es grupal o individual (dependiendo de las características específicas de cada caso) donde se guía al estudiante hacia la carrera que más se identifique con su vocación.

Los procesos de orientación vocacional llevados a cabo tienen como fin generar un espacio de encuentro donde el estudiante pueda plantear sus dudas, incertidumbres, pensamientos y deseos. La duración estimada del proceso es de ocho encuentros (alrededor de un mes), con una frecuencia de dos reuniones semanales, de una hora de duración cada una. Están dirigidos a estudiantes del último año del Nivel Secundario (tanto de la capital, como del interior provincial) o para adultos que decidan iniciar o modificar un proyecto de vida asociado a la elección de una carrera. Se cuenta con profesionales psicólogas que brindan asesoramiento, por medio de charlas informativas sobre la vida universitaria, acompañando en visitas guiadas a Facultades y Escuelas de la UNCA. Además, el equipo de trabajo de la Dirección de Orientación Vocacional participa en el Programa la UNCA+Cerca (promovido por la Secretaria de Extensión Universitaria), también brindan charlas informativas en escuelas que requieren dicha participación, pensando en la matrícula potencial de la universidad y una inserción institucional efectiva.

El presente proyecto surge con la idea de brindarle a la Dirección de Orientación Vocacional dependiente de la Secretaria de Bienestar Universitario y Asuntos Estudiantiles de la UNCA una aplicación móvil que permita realizar en forma automatizada el test de orientación vocacional, el cual se realiza actualmente mediante formularios impresos. La aplicación tiene como finalidad servir como un instrumento, para que el estudiante tenga una orientación de cuáles son sus áreas ocupacionales de preferencia a la hora de optar por una carrera y proveerle información sobre las carreras que se imparten en la UNCA relacionadas a dichas áreas. Cabe destacar que esta aplicación es solo un primer paso en la búsqueda vocacional, por lo cual se aconsejará al usuario (estudiante) participar del Proceso de Orientación Vocacional que brinda la Dirección de Orientación Vocacional.

La aplicación móvil además permite al usuario:

- Obtener información sobre las diferentes actividades que se realizan en la Dirección de Orientación Vocacional.
- Obtener información sobre los distintos programas de becas y otros servicios que

ofrece la Secretaria de Bienestar Universitario y Asuntos Estudiantiles.

- Georeferenciar en un mapa las localizaciones de las distintas unidades de la UNCA.

Mediante la aplicación móvil se recaba información del usuario (edad, procedencia, escuela secundaria, sexo, e-mail, fecha, área y resultado del test) y se almacena en un servidor para posteriormente ser procesada para la obtención de estadísticas que sean de utilidad para la UNCA.

Líneas de Investigación, Desarrollo e Innovación

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

- Procesos de orientación vocacional. Test de orientación vocacional existentes.
- Factores que influyen en la deserción estudiantil y políticas universitarias tendientes a reducir los índices de deserción.
- Metodologías para el desarrollo de aplicaciones móviles.

Para la ejecución del proyecto de I+D se planificaron las siguientes 7 etapas:

1. Estado del arte: mediante el estudio de bibliografía actualizada y la asistencia a Cursos, Congresos y Workshops específicos, se profundizará en el estado del arte de los procesos de orientación vocacional, factores que influyen en la deserción estudiantil, y políticas universitarias tendientes a reducir los índices de deserción.
2. Estudio, análisis y selección de diferentes test de orientación vocacional, junto a las psicólogas dependientes de la Dirección de Orientación Vocacional.
3. Estudio, análisis y selección de metodologías para el desarrollo de aplicaciones móviles.
4. Relevamiento de los requerimientos para la aplicación móvil y adaptación del test seleccionado a las particularidades de los estudiantes catamarqueños del Nivel Secundario.
5. Desarrollo de la aplicación móvil.
6. Análisis de los resultados obtenidos.

Resultados obtenidos/esperados

La aplicación móvil se desarrolló siguiendo los lineamientos de la metodología propuesta por Gasca Mantilla et. al. [14] para el desarrollo de aplicaciones móviles.

Actualmente, la aplicación se encuentra en la fase de pruebas (correspondiente a la etapa 5) con el objetivo de verificar el funcionamiento de la aplicación en diferentes escenarios y condiciones. En esta fase se están realizando 1) pruebas simulando el escenario y emulando el dispositivo móvil, explorando todas las utilidades y funciones de la aplicación; 2) pruebas de campo en equipos reales para medir el desempeño y el rendimiento de la aplicación; y 3) análisis de las 6 M's para valorar el potencial de éxito del servicio. Para esto último, profesionales y estudiantes avanzados de la carrera de Ingeniería en Informática, utilizan el método de evaluación de las 6 M's, y califican la presencia de los seis atributos en la aplicación desarrollada. Los atributos analizados son movimiento (*movement*), momento (*moment*), yo (me), método (*method*), dinero (*money*) y máquinas (*machines*).

Se espera que el uso de esta aplicación permita a la UNCA:

- Masificar la cantidad de estudiantes que puedan acceder al test, sobre todo de aquellos estudiantes que viven en el interior de la provincia y no pueden concurrir a la Dirección de Orientación Vocacional a realizar el test.
- Ayudar a reducir la deserción estudiantil, ya que diversos estudios han demostrado que un alumno que eligió su carrera considerando sus habilidades, gustos e intereses (información que se obtiene a

través de los test de orientación vocacional), tendrá mayores posibilidades de éxito escolar.

- Delinear líneas de trabajo y de articulación anticipando la población de potenciales ingresantes, a partir de los resultados que se obtengan de los test.

Formación de Recursos Humanos

La línea de I+D presentada cuenta con una Doctora en Ingeniería (mención Sistemas de Información), dos Magister en Ingeniería de Software, una Especialista en Educación y Nuevas Tecnologías, dos Licenciados en Sistemas de Información y dos alumnos avanzados de Ingeniería en Informática. Los profesionales del equipo son docentes investigadores Categoría III y IV.

De los estudiantes que forman parte del equipo, uno de ellos se encuentra finalizando su Trabajo Final de Carrera, cuyo objetivo principal es el desarrollo de la aplicación móvil para el test de orientación vocacional.

También 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.
- Participación en talleres o workshops de herramientas informáticas relacionadas con la temática
- Convocar e introducir a los alumnos de la carrera de Ingeniería en Informática a la realización de actividades de investigación y desarrollo.

Referencias

[1] González Fiegehen, L. E., & Espinoza Díaz, O. (2008). *Deserción en educación superior en América Latina y el Caribe*. Revista Paideia, 45, 33-46. Disponible en: https://www.researchgate.net/publication/275275484_Desercion_en_educacion_superior_en_America_Latina_y_el_Caribe_2008-16

[2] Sánchez Amaya, G., Navarro Salcedo, W., & García Valencia, A. (2009). *Factores de deserción estudiantil en la Universidad Surcolombiana*. Paideia Surcolombiana, 1(14), 97-103. Disponible en: https://www.journalusco.edu.co/index.php/paideia/article/view/1083/2108

[3] Ciano, N., Castignani, M. L., & García, M. N. (2011). *Exploración del abandono universitario en estudiantes de las carreras de la Facultad de Psicología de la Universidad Nacional de La Plata*. 3er Congreso Internacional de Investigación. Universidad Nacional de La Plata. Facultad de Psicología. Disponible en: http://www.memoria.fahce.unlp.edu.ar/trab_eventos/ev.1516/ev.1516.pdf

[4] Santoyo, M. K. F., & Patiño, J. C. S. (2017). *Motivos de deserción de estudiantes de Licenciatura durante su primer año cursado, en el Instituto Tecnológico Superior de Irapuato. Factores asociados al abandono. Tipos y perfiles de abandono.* VII Conferencia Latinoamericana sobre el Abandono en la Educación Superior (CLABES). Disponible en: http://revistas.utp.ac.pa/index.php/clabes/article/view/1626/2363

[5] Huesca Ramírez, M.G., & Castaño Corvo, M. B. (2007). *Causas de deserción de alumnos de primeros semestres de una universidad privada*. Revista Mexicana de Orientación Educativa, 5(12), 34-39. Disponible en: http://www.alfaguia.org/alfaguia/files/1319582164causas%20de%20desercion%20en%20una%20universidad%20privada.pdf

[6] Escobar, D.J. (2013). *Elección profesional y deserción universitaria. Del redireccionamiento del plan de vida*. Revista Electrónica Psyconex: Psicología, Psicoanálisis y Conexiones. Vol. 5 (8). 2013. ISSN 2145-437X. Disponible en:

http://aprendeenlinea.udea.edu.co/revistas/index.php/Psyconex/article/viewFile/20983/17583

[7] Infante Tavío, N. I., Mendo Alcolea, N., & Vázquez Sánchez, M. (2012). *Factores determinantes de la deserción escolar en el Policlínico Docente" Frank País García"*. Medisan, 16(4), 526-531. Disponible en: http://scielo.sld.cu/pdf/san/v16n4/san06412.pdf

[8] Betancourth Sánchez, L. J. (2016). *Orientación vocacional y profesional en la juventud colombiana*. Tesis para optar al título de Especialista en Docencia Universitaria. Universidad Militar Nueva Granada. Colombia. Disponible en: https://repository.unimilitar.edu.co/bitstream/handle/10654/14245/BetancourthS%C3%A1nchezLinaJhulieth.pdf?sequence=1&isAllowed=y

[9] Margain Pérez, A.K. & Murillo García, F. (2014). *La orientación vocacional como estrategia para favorecer el éxito escolar*. III Seminario del Sistema de Información de Estudiantes, Egresados y Empleadores (SIEEE). Disponible en: https://www.azc.uam.mx/sieee/cuartoseminario/ponencias/ponencia06.pdf

[10] De León Mendoza, T. & Rodríguez Martínez, R. (2008). *El efecto de la orientación vocacional en la elección de carrera*. Revista Mexicana de Orientación Educativa, 5(13), 10-16. Disponible en: http://pepsic.bvsalud.org/pdf/remo/v5n13/v5n13a04.pdf

[11] Albeza Martin, M. A. & Mendoza Pontiffe, L. (2016). *Vocación y Permanencia Educativa*. IV Encuentro Nacional de Servicios de Orientación Universitaria "Hacia una Mirada Interdisciplinaria". Disponible en: http://blogs.unlp.edu.ar/encuentroorientacionuniversitaria/files/2016/11/ponencias-orientacion-vocacional.pdf

[12] Rodríguez Vásquez A. & Ochoa Ariza, M.A. (2012). *Relación entre el proceso de Orientación Vocacional implementado en las Instituciones de Educación Media del Departamento del Cesar y la deserción estudiantil en la Universidad Popular del Cesar*. Tesis para optar al Título de Magister en Educación. Sistema de Universidades Estatales del Caribe Colombiano. Disponible en: https://dialnet.unirioja.es/descarga/articulo/6320329.pdf

[13] Gasca Mantilla, Maira Cecilia, Camargo Ariza, Luis Leonardo, & Medina Delgado, Byron. (2014). *Metodología para el desarrollo de aplicaciones móviles*. Tecnura, 18(40), 20-35. Disponible en: http://www.scielo.org.co/pdf/tecn/v18n40/v18n40a03.pdf

Aplicaciones interactivas usando realidad aumentada y realidad virtual

Claudia Banchoff, Laura Fava, 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
{cbanchoff, lfava, ales}@info.unlp.edu.ar, smartin@linti.unlp.edu.ar

RESUMEN

Este artículo describe la línea de investigación, desarrollo e innovación vinculada al desarrollo de aplicaciones usando Realidad Aumentada (RA) y Realidad Virtual (RV) 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 educativas interactivas, materiales didácticos, objetos de aprendizaje, así como también, aplicaciones destinadas a la recreación de niños durante tratamientos terapéuticos y/o de rehabilitación prolongados o para acompañar las visitas en edificios históricos.

En este artículo, se continúa con las líneas de I+D presentadas en WICC 2018 (Díaz, J., Banchoff, C., Fava, L., Schiavoni, A., Martin, S. 2018), donde se analizaron y evaluaron entornos de desarrollo para aplicaciones de gamificación usando RA y RV. A partir de esta evaluación, se conocen las capacidades y limitaciones de los principales entornos de desarrollos, pudiendo decidir cuál es la mejor opción para cada tipo de aplicación. En esta etapa se hará enfoque en el desarrollo de aplicaciones usando RA y RV destinadas a escuelas secundarias, al Museo de Ciencias Naturales de la Universidad Nacional de La Plata y al Hospital de Niños de La Plata.

Palabras clave: juegos serios, aprendizaje basado en juegos, realidad aumentada, realidad virtual, gamificación.

CONTEXTO

La línea de investigación centrada en el desarrollo y aplicación de juegos serios se viene trabajando en el LINTI desde hace una década. Los desarrollos y experiencias realizados están vinculados a distintas problemáticas sociales. En el último año se ha focalizado en la incorporación de tecnologías de RA/RV como un elemento innovador y motivador para complementar las actividades que se trabajan en el aula de la escuela.

En este sentido se ha consolidado el uso de videojuegos en el aula tales como RAÍCES (Fava, L., Banchoff, C., Nomdedeu, L., Martin, S., 2017). y realizado experiencias de RA utilizando herramientas como Scratch con niños y niñas de primaria en dos jornadas realizadas en la Facultad de Informática. Estas actividades dan el marco propicio para el desarrollo conjunto de proyectos de extensión e investigación aplicada, en los cuales se trabaja con distintas instituciones educativas.

La línea de investigación que se presenta en este trabajo, "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 .

1. 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 que incorporan aspectos de realidad aumentada (RA) y/o realidad virtual (RV) son instrumentos muy poderosos para ser usados en múltiples contextos, en especial, en el ámbito educativo.

Existen muchas experiencias con distintos niveles de aceptación. En nuestra región el desafío es generar actividades que puedan ser llevadas adelante en los establecimientos de gestión pública que sólo cuentan, en el mejor de los casos, con netbooks entregadas en el marco de distintos programas gubernamentales, pero que no tienen acceso a dispositivos móviles y cascos de RV.

El Observatorio de Innovación Educativa del Tecnológico de Monterrey, en su último reporte, analiza el rol que juega en el ámbito educativo la inclusión en los procesos de enseñanza-aprendizaje de tecnologías de RA y RV. (Edu Trends 2017). Si bien los datos son interesantes, en nuestra región no pueden aplicarse los mismos resultados. No hay duda, que estas tecnologías pueden resultar altamente atrayentes a docentes y estudiantes, pero no se cuenta con estudios y resultados registrados en nuestra región que sirvan de modelo que retroalimenten nuevas experiencias.

Hay muchas aplicaciones disponibles, pero en la gran mayoría de los casos, las mismas son utilizadas por docentes “curiosos”. En muchos casos se trata de aplicaciones en idiomas extranjeros o con licencias privativas que limitan su uso.

En base a experiencias previas, se trabaja en nuevos desarrollos realizados en conjunto con lo equipos docentes que trabajan en los distintos establecimientos con los que se trabaja.

La aplicación de tecnologías de RA en el ámbito de las escuelas es, desde el punto de vista de acceso a los recursos físicos necesarios, más simple que la incorporación de RV. Es común encontrar acceso a dispositivos móviles, tanto en los equipos docentes como en los mismos estudiantes. El uso de RA, a través de dispositivos móviles ha mostrado nuevas posibilidades e implicancias para diferentes audiencias y contextos (Das P., Zhu Mo, McLaughlin, L., Bilgrami, Z., Milanaik, R., 2017). Sin embargo, para utilizar aplicaciones de RV, las soluciones comerciales que podemos encontrar son los Oculus Rift, HTC Vive o Sony PlayStation RV, las cuales ofrecen una experiencia bastante lograda, pero requieren el apoyo de un PC muy potente que no está al alcance de de la comunidad educativa. Afortunadamente, también existen opciones más económicas, como Samsung Gear VR, o, incluso, las Cardboard de Google, que funcionan con un celular pero con una experiencia menos logrado. Estos recursos limitan notablemente el tipo de desarrollo y actividad que se puede realizar. En el ámbito de la escuela, es importante que los desarrollos realizados puedan ser usados con recursos de bajo costo.

Para el desarrollo de aplicaciones, tanto de RA como RV, hay un gran número de opciones disponibles que van desde opciones privativas a libres, y que ofrecen mayor o menor compatibilidad con plataformas. Algunas de ellas son Unreal Engine 4 de Epic Games, Unity de Unity

Technologies, y CryEngine de Crytek. Algunas de éstas proveen licencias que habilitan su uso en ámbitos académicos y, si bien tienen restricciones respecto al sistema operativo del editor, es posible desarrollar aplicaciones para casi todas las plataformas disponibles.

Por lo analizado se puede concluir que el uso de técnicas de RA brinda un mayor potencial a las aplicaciones. En el caso de la 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 basadas 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. En otros contextos, por ejemplo en edificios históricos, el uso de RA permite integrar objetos del mundo real con animaciones e información adicional, favoreciendo su difusión, y preservando el patrimonio cultural tangible e intangible (Shengli Xu, 2018)

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Los ejes en los que se está investigando están relacionados con el desarrollo y uso de videojuegos y aplicaciones interactivas usando RA y RV en ámbitos educativos de nivel inicial y primario, y en otros ámbitos relacionados con edificios históricos y de recreación en general.

A continuación se detallan los ejes de investigación:

- Utilización de las herramientas analizadas en el desarrollo de aplicaciones en los diferentes ámbitos.
- Estudio y análisis de nuevas versiones de los softwares específicos con el fin de aprovechar las funcionalidades provistas.
- Análisis de los contextos reales en los cuales se van a desarrollar las aplicaciones planteadas.
- Extensión de los prototipos desarrollados como base para la implementación de los productos finales.
- 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.

3. RESULTADOS Y OBJETIVOS

El objetivo general de esa línea de trabajo es aplicar técnicas de RA y RV en desarrollos concretos, en especial, para el ámbito educativo.

Para cumplir con el objetivo general, se proponen los siguientes objetivos específicos:

- Analizar los distintos contextos de uso de las aplicaciones, los cuales direccionarán las tecnologías a utilizar.
- Desarrollar aplicaciones interactivas que puedan utilizarse en el ámbito de la escuela, el museo de Ciencias

Naturales y el Hospital de Niños de La Plata.

- Construir casos de prueba de las aplicaciones desarrolladas con el fin de aplicarlos en los entornos para los cuales fueron desarrolladas y comprobar su adecuación y usabilidad.
- Continuar promoviendo esta temática en el marco del desarrollo de las tesinas de grado.

En el marco de esta línea de investigación se están trabajando en los siguientes desarrollos:

Sinfonía Terrestre, es una aplicación desarrollada para el Museo de Ciencias Naturales de La Plata, en la cual se aplican técnicas de RA en el recorrido de tres (3) salas del museo. La Fig. 1 muestra una captura de la interfaz de usuario.de la aplicación Sinfonía Teerestre.

Fig. 1: Diseño de la aplicaciones para el Museo

ERA, es una aplicación que, en principio, será destinada al Hospital de Niños de La Plata, y que surge como un paratexto de un libro desarrollado en el marco del proyecto "Cuente con nosotros" llevado a cabo por la cátedra Lenguaje Visual de la Facultad de Bellas Artes de la UNLP. A través de técnicas de RA se extenderán libros, en este caso, KOTA CORTA, con animaciones y juegos sencillos que será destinado a niños y niñas que se encuentran realizando tratamientos prolongados en el hospital.

Fig. 2: Personaje del libro Kota Corta

Se prevé finalizar los desarrollos y realizar las pruebas correspondientes en las instituciones para las cuales fueron desarrolladas.

En esta misma línea se comenzó a trabajar en aplicaciones, que utilizan técnicas de RV, destinadas a estudiantes de nivel primario. La temática y especificación de las mismas está siendo consensuada con docentes de tres escuelas de la región.

Los trabajos presentados en este artículo se llevan a cabo de manera interdisciplinaria entre docentes y alumnos de la Facultad de Informática y directivos del Museo de Ciencias Naturales. Asimismo, se mantienen relaciones con docentes de la Universidad Politécnica de Valencia con el objetivo de desarrollar proyectos en forma conjunta y consolidar la relación bilateral.

4. 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 y un equipo de docentes de la cátedra Diseño multimedial de la Facultad de Bellas Artes 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, un desarrollo para el Hospital de niños y aplicaciones para escuelas primarias.

5. BIBLIOGRAFÍA

Das, P.; Zhu, M.; McLaughlin, L.; Bilgrami, Z.; Milanaik, R.L. (2017). *Augmented reality video games: new possibilities and implications for children and adolescents*. Multimodal Technologies and Interaction, 2017; 1:8.

Díaz, J., Banchoff C, Fava, L., Schiavoni, A., Martin S. (2018). *Juegos serios y aplicaciones interactivas usando realidad aumentada y realidad virtual*. XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), ISBN: 978-987-3619-27-4. Abril, 2018.

Edu Trends, (2017). *Observatorio de Innovaciońń Educativa del Tecnológico de Monterrey*. Disponible en https://goo.gl/zJFx8L. Último acceso, marzo de 2019.

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.

Shengli Xu (2018). *Intangible Cultural Heritage Development Based on Augmented Reality Technology*. 2018 International Conference on Robots & Intelligent System (ICRIS). DOI: 10.1109/ICRIS.2018.00094

Aprendizaje mediado por Tecnología Móvil

Mg. Roberto Bertone[1], Mg. José Luis Filippi[2], Lic. Guillermo Lafuente[3], Mg. Carlos Ballesteros[4], Lic. Gustavo Lafuente[5], I.S. Daniel Perez[6], I.S. Sofía Aguirre[7], A.S. Alejandra Mansilla[8]
LIAU[9] - Facultad de Ingeniería – UNLPam.
pbertone@ada.info.unlp.edu.ar[1]
{filippij[2], lafuente[3], balleste[4], gustavo[5], perezd[6], aguirres[7], mansilla[8]}@ing.unlpam.edu.ar
[9]Laboratorio de Investigación de Ambientes Ubicuos

Resumen

La integración de las TICs en las instituciones educativas es un proceso que se desarrolla desde hace varias décadas en la República Argentina. Si bien existe un fuerte consenso en la universalización del acceso a las diferentes tecnologías de la información y las comunicaciones, la discusión vigente reside en la incorporación de los dispositivos móviles (teléfonos de última generación) como instrumento mediador del Proceso de Enseñanza-Aprendizaje.

En la actualidad la educación no se encuentra limitada a los entornos formales conformado por las instituciones educativas en sus diferentes niveles, se integra con procesos formativos que tienen lugar a través de los nuevos dispositivos tecnológicos de última generación, caracterizados por su miniaturización, movilidad y conectividad permanente; haciendo posible el acceso a múltiples contextos de aprendizaje virtuales. Resulta entonces imprescindible estudiar las tendencias del aprendizaje móvil en el mundo desarrollado, realizar las adaptaciones necesarias e implementarlas en el ámbito educativo.

El propósito del proyecto, de naturaleza teórico-práctico, es el de indagar las posibilidades que ofrecen los dispositivos móviles (teléfonos inteligentes, tabletas digitales) como instrumentos aplicados a diferentes situaciones de enseñanza y aprendizaje.

Bajo esa premisa se intenta llevar a cabo un trabajo experimental en la Facultad de Ingeniería - UNLPam, desarrollando un escenario de aplicación real con la utilización de éste tipo de dispositivos.

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

Con más de 5900 millones de usuarios de telefonía móvil en el mundo, los dispositivos de última generación han transformado nuestra manera de vivir. Aunque en todos los sectores de la sociedad se hace uso intensivo de esta tecnología, los educadores y los responsables de formular las políticas educativas no han aprovechado su potencial para mejorar el proceso educativo. **[1]**

La aparición continua de nuevas herramientas tecnológicas móviles avanza en lapsos cortos de tiempo de forma sorprendente. Por ello es importante que los educadores comprendan estas innovaciones, sería ideal que la educación y la tecnología evolucionaran en paralelo y que las necesidades educativas

promovieran el progreso tecnológico además de adaptarse a él.

Múltiples necesidades se presentan en el ámbito educativo en sus diferentes niveles entre los cuales se pueden mencionar: desarrollo de contenidos bibliográfico digital de producción local y regional, incentivar el uso de aplicaciones móviles que posibiliten el intercambio de información entre los actores que conforman la comunidad educativa (videoconferencias entre docentes, estudiantes, directivos, y comunidad en general), incorporar el uso de plataformas de formación virtual abiertas, que posibiliten incorporar recursos y/o materiales didácticos en diferentes formatos, mantener un canal de comunicación fluido entre la institución educativa y el contexto local, regional, y nacional, e implantar estrategias metodológicas para alcanzar un aprendizaje cooperativo y colaborativo.

Mark Weiser **[2]** 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 aplicaciones tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las necesidades de los usuarios que transitan en el ámbito educativo, profundizando en el aprendizaje móvil.

Situación Actual del Problema

La aparición de dispositivos tecnológicos caracterizados por su miniaturización, gran capacidad de almacenamiento y gran velocidad de procesamiento, dieron origen a nuevos entornos en la comunicación de las personas, con una hipercomunicación caracterizada por la multimedialidad. El envío de mensajes textuales seguidos de imágenes, sonidos y videos es parte de la actividad cotidiana.

La educación influenciada por ésta realidad, acepta la necesidad de reconfigurar el proceso educativo, partiendo de algunas premisas:

a. Redefinir nuevas acciones para docentes y estudiantes.

El creciente número de dispositivos móviles y la conectividad a internet de forma permanente, permite a los estudiantes adquirir nuevos contenidos en línea y configurar sus conocimientos a partir de intereses propios. El estudiante es artífice de su futuro, asume un papel activo, decide que aprender y cómo.

El docente deja de ser el poseedor indiscutible del conocimiento para cumplir un rol diferente, facilitador del aprendizaje, brindando ayuda al estudiante a beneficiarse con la enorme cantidad de información que las tecnologías móviles ponen a su alcance.

Frente a la realidad del mundo digital que rodea a los estudiantes de hoy en día, es justo que los docentes replanteen su rol en la construcción del conocimiento sobre ellos. De allí que la competencia exigida a un profesor del Siglo XXI es preciso que se enfoquen "en las necesidades de los alumnos, supervisando su búsqueda de información e intentando facilitar la búsqueda de información individual de los alumnos ya que el papel de suministrador de conocimiento ha sido superado por las TIC" [3].

Mucho se habla de cuáles serían las competencias digitales para un docente, un acercamiento a ello lo propone Núñez-Torrón Stock [4], quien plantea cinco competencias básicas para el docente en la educación, y entre ellas se pueden destacar dos:

1. Usar herramientas de trabajo en línea.
2. Utilizar dispositivos móviles en el aula.

En la mayoría de las universidades de todo el mundo se observa la incorporación de los sistemas de Gestión de Aprendizaje (LMS), para implementar entornos virtuales que permiten la gestión del aprendizaje. **[5]** Sin embargo algunas investigaciones indican que

las motivaciones de los estudiantes son vistas como una obligación al participar en entornos de aprendizaje institucionales cerrados, mientras que la motivación es mayor cuando el proceso formativo es abierto a diferentes tecnologías en contextos informales. Estos espacios son conocidos como PLE (Personal Learning Environments) donde los estudiantes configuran su propio entorno. **[6]**
Algunos autores ya incorporan un nuevo concepto, el de mPLE (Mobile Personal Learning Environments) como una nueva estructura de aprendizaje para la generación que hace uso de los dispositivos móviles. **[7]**

b. Continuidad del aprendizaje.

El aprendizaje móvil permitirá a los estudiantes dar continuidad a su formación a lo largo de toda su vida, a partir de plataformas que posibilitan el aprendizaje móvil caracterizado por:

- Poseer gran cantidad de información.
- Ser omnipresentes.
- Disponer de gran diversidad de recursos materiales.
- Fomentar la participación colaborativa.

Actualmente los estudiantes y futuros profesionales acceden a la información a través del aprendizaje informal, han dejado de ser únicamente consumidores para ser productores de información.
En el caso particular de los estudiantes universitarios no solo deben dominar las disciplinas correspondientes a su carrera, además deben incorporar habilidades transversales como el pensamiento crítico, resolución de problemas, persistencia y trabajo colaborativo. Sin embargo, en muchos países no se están desarrollando estas habilidades. **[8]**

c. Grandes volúmenes de datos. Big data.

Por Big data entendemos al tratamiento y análisis de enormes repositorios de datos, tan grandes y desestructurados que resulta imposible tratarlos con las herramientas de base de datos y analíticas convencionales. Permite procesar grandes volúmenes de datos a través de numerosas computadoras pequeñas unidas en redes de procesamiento.
Disponer y aprovechar ésta tecnología es una obligación para todas aquellas instituciones que manipulan grandes volúmenes de datos, en nuestro caso particular la universidad.
La gran cantidad de información accesible a partir de variadas herramientas conlleva una serie de consideraciones éticas relacionadas a la propiedad de los datos y la privacidad.
Los investigadores que estudian el aprendizaje en línea, los sistemas de tutoría inteligente, los laboratorios virtuales, las simulaciones y los sistemas de gestión del aprendizaje están explorando maneras de entender y utilizar mejor la analítica del aprendizaje, a fin de mejorar la actividad del docente y seguir avanzando en la educación para todos. **[9]**
En éste contexto el punto de partida debe ser siempre pedagógico. No se trata de innovar por innovar. Hay que explotar el potencial que ofrecen las nuevas herramientas tecnológicas que hacen posible el aprendizaje móvil, a partir de una planificación docente previa dentro de un marco teórico formalizado.
La herramienta en sí misma no ofrece por si sola resultados, pero un buen docente sabrá aprovechar las ventajas que brinda para un aprendizaje móvil de calidad.
Por ello el aprendizaje móvil (mLearning) puede incorporarse como refuerzo del aprendizaje formal, ampliando la oferta educativa y la modalidad. Lo importante es innovar en el proceso educativo en su conjunto y no solo con el dispositivo tecnológico.
El entorno educativo cambia **[10]**, la educación se presenta como la formación de

los educandos en competencias, destrezas, habilidades para desempeñarse en un nuevo espacio social, el digital **[11]**. Nuevos escenarios educativos, los mismos actores con un nuevo rol, y la implantación de las tecnologías móviles es el desafío actual **[12].**

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.

- Identificar el porcentaje de inserción de dispositivos móviles en la institución (teléfonos y/o tablets).
- Analizar la inserción del aprendizaje móvil en las universidades de todo el mundo.
- Estudiar como el aprendizaje móvil puede cerrar la brecha entre el aprendizaje formal y el informal.
- Identificar los escenarios educativos que muestren aspectos móviles en las actividades formativas del ámbito académico propio.
- Examinar herramientas tecnológicas móviles orientadas al proceso educativo de acceso libre y gratuito.

Segundo y Tercer año.

- Definir el enfoque pedagógico que posibilite un aprendizaje móvil de calidad.
- Estudiar las tecnologías disponibles para implementar acciones de aprendizaje móvil en el contexto educativo universitario.
- Definir prioridades según necesidades de inmediatez en la virtualización de las disciplinas que se han de impartir.
- Gestionar el uso de plataformas de formación virtual que posibilite una doble modalidad, presencial y mediada por las nuevas tecnologías móviles.
- Incorporar el uso de redes sociales que favorezcan la práctica educativa.
- Confeccionar objetos de aprendizaje acorde a las herramientas disponibles en la nube y a las características de los dispositivos móviles que van a operar el producto final.
- Desarrollar aplicaciones móviles a partir de las necesidades que se presenten durante el transcurso del proceso educativo virtual.
- Capacitar a la comunidad educativa en general en el desarrollo de objetos de aprendizaje a partir de los requerimientos de carácter institucional.
- Difundir los avances a toda la comunidad universitaria los progresos a medida que van trascurriendo, a través de jornadas, congresos y/o revistas científicas.
- Propiciar el intercambio de información permanente con grupos de investigación que den valor agregado a nuestra actividad profesional.

Cuarto año.

- Instituir las aplicaciones desarrolladas para dispositivos móviles que hayan alcanzado buen nivel de aceptación.
- Conformar un repositorio de objetos de aprendizajes de acceso libre.
- Registrar y difundir los resultados alcanzados con la finalidad de que se puedan utilizar, ampliar y mejorar a través de trabajos futuros.
- Disponer de un catálogo de trabajos realizados en el sitio web del grupo de investigación GIAU (Grupo de Investigación de Ambientes Ubicuos) a partir del cual los interesados podrán acceder al material requerido.

Presentar en jornadas, congresos y/o revistas de todo el mundo los resultados alcanzados.

Resultados Obtenidos/Esperados

El proyecto da inicio a su actividad de investigación y desarrollo durante el 2019, por lo que sus integrantes se encuentran efectuando las tareas enmarcadas en el primer año.

Al finalizar el proyecto se espera contribuir en la incorporación de aplicaciones tecnológicas móviles a través de las cuales se puedan ofrecer servicios que satisfagan las

necesidades de los usuarios que transitan en el ámbito educativo de la UNLPam y la UNLP, profundizando en el aprendizaje móvil.

Formación de Recursos Humanos

Director de Proyecto
Co-Director de Proyecto
6 Investigadores

Referencias

[1] Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2013). El futuro del aprendizaje móvil. Implicaciones para la Planificación y la Formulación de Políticas. http://unesdoc.unesco.org/images/0021/002176/217638s.pdf

[2] Weiser M. (1991), The Computer for the Twenty-First Century.

[3] webdelmaestrocmf.com, "Estas son las capacidades digitales de un docente moderno." [Online]. Available: http://webdelmaestrocmf.com/portal/5-competencias-digitales-que-deben-tener-los-profesores-actuales/?utm_source=blogsterapp&utm_medium=Facebook. [Accessed: 27-Oct-2017].

[4] A. Núñez-Torrón Stock, "5 competencias digitales que deben tener los profesores actuales." [Online]. Available: http://www.ticbeat.com/educacion/5-competencias-digitales-profesores/. [Accessed: 27-Oct-2017].

[5] M. P. Prendes, "Plataformas de campus virtual con herramientas de software libre: Análisis comparativo de la situación actual en las universidades españolas," Murcia, Project Report EA-2008-0257, Jun. 2009. [Online]. Available: http://www.um.es/campusvirtuales/informe_final_CVSL_SF.pdf

[6] F. J. García-Peñalvo, "Docencia," in Libro blanco univ. digital 2010, J. Laviña Orueta and L. Mengual Pavón, Eds. Barcelona, Spain: Ariel, 2010, pp. 29-61.

[7] G. Attwell, J. Cook, and A. Ravenscroft, "Appropriating Technologies for Contextual Knowledge: Mobile Personal Learning Environments," in Best Practices for the Knowledge Society. Knowledge, Learning, Development and Technology for All, M. D. Lytras, P. O. de Pablos, E. Damiani, D. Avison, A. Naeve, and D. G. Horner, Eds. Springer Berlin Heidelberg, 2009, pp. 15-25.

[8] (World Economic Forum, 2015).

[9] Bienkowski, M., Feng, M., & Means, B. (2012). Enhancing teaching and learning through educational data mining and learning analytics: An issue brief.US Department of Education, Office of Educational Technology, 1-57.

[10] Weiser, M. (1998) *The future of Ubiquitous Computing on Campus*. Communications of ACM, 41-1, January 1998, 41-42.

[11] Malani R,. Griswold W, Simon B, (2009) *Public Digital Note-Taking in Lectures*. Ubicomp 2009.

[12] Richards M, Woodthorpe J, (2009), *Introducing TU100 "My Digital Life": Ubiquitous computing in a distance learning environment*. Ubicomp 2009.

Desarrollo de entornos accesibles para la Educación Superior en Ingeniería: Aplicación a la carrera de Ingeniería en Sistemas de Información

Alejandra Jewsbury, María Laura Destefanis, Fanny Montoya, Elizabeth Jeinson, Inés Haefeli, Alejandra Odetti, Silvia Arias
Dep. Ingeniería en Sistemas de Información/ Facultad Regional Córdoba/ Universidad Tecnológica Nacional
Maestro López esq. Cruz Roja (Ciudad Universitaria, Córdoba)
ajewsbury@gmail.com, mldestafanis@gmail.com, femlys@gmail.com, ejeinson@gmail.com, ineshaefeli@gmail.com, maaodetti@gmail.com, sarias@iua.edu.ar

Resumen

Un mundo accesible para todos y todas y en cualquier momento y lugar es el anhelo de la sociedad actual. Las minorías no escuchadas y por años silenciadas alzan hoy la voz para acceder como iguales a los beneficios de la vida en sociedad. Desde distintas esferas de la sociedad civil y de las instituciones políticas emergen acciones que permitan a las personas con discapacidad acceder a los beneficios de la educación y para los estudios universitarios en la Argentina se encuentra con una ley de alcance nacional. Es así que la universidad, y en especial las carreras de ingeniería deben observarse y modificar sus aspectos formativos, sus didácticas y sus espacios edilicios para hacer posible una universidad inclusiva. Este proyecto analiza la enseñanza de la ingeniería desde la visión de la accesibilidad y propone modelos de enseñanza y recursos tecnológicos que permitan el acceso, la permanencia y el desarrollo pleno de personas con discapacidad.

Palabras clave: accesibilidad, didáctica, diseño curricular

Contexto

El PID en el que se basa esta presentación es el identificado como TEUTCO0004981.

Tema prioritario del Programa de Tecnología Educativa y Enseñanza de la Ingeniería en que se inserta: La enseñanza de la ingeniería y la formación de los ingenieros.

Introducción

La accesibilidad es definida como (Peralta Morales, 2007) [1] “el conjunto de características que debe disponer un entorno, producto o servicio para ser utilizable en condiciones de confort, seguridad e igualdad por todas las personas y, en particular, por aquellas que tienen alguna discapacidad". En el marco de la educación superior se observan en el concepto de accesibilidad tres aspectos (López, 2015), el primero referido al entorno presencial (universidad, aula, aula de laboratorio) o virtual (plataforma de aprendizaje, aula virtual, redes sociales), el segundo referido al producto (asimilable a los recursos educativos seleccionados por los docentes, sean en

formato físico o digital) y el tercero asociado a los servicios educativos (las prácticas docentes de enseñanza, incluyendo las estrategias e instrumentos de evaluación). En referencia a la presencia de plataforma de educación virtual se encuentra mucha bibliografía e importantes portales con investigaciones y producciones. Se destaca el trabajo realizado en el marco del Proyecto de Educación Superior Virtual Inclusiva en América Latina (ESVIAL), enmarcado en el Programa Alfa# de la Unión Europea, que tiene por objeto mejorar la educación virtual en América Latina a través de la creación de metodologías que establecen un conjunto de estándares de accesibilidad para la incorporación de personas con discapacidad a los estudios universitarios. Este proyecto finalizado en el año 2015 coloca a disposición de académicos, investigadores y gestores universitarios una Guía Metodológica. El modelo de Acreditación ESVIAL es abierto y está disponible en www.esvial.org. La inserción y difusión de la accesibilidad en estudios universitarios en Argentina se encuentra enmarcada en la Ley de Educación Superior Nro. 24521 sancionada en el año 1995 y la modificación del año 2002, Ley 25573 [2]. El Artículo 1 de la Ley 24521 establece que es el Estado el que tiene la responsabilidad indelegable de la prestación del servicio de educación superior de carácter público y reconoce y garantiza el derecho a cumplir con ese nivel de la enseñanza a todos aquellos que quieran hacerlo y cuenten con la formación y capacidad requeridas.

En la modificación del año 2002 se incorpora, en el artículo 1, la declaración de que es el Estado el garante de la accesibilidad al medio físico, servicios de interpretación y los apoyos técnicos necesarios y suficientes, para las personas con discapacidad. La modificación del Artículo 2 de la Ley 25573 agrega el inciso f) que expone “Las personas con discapacidad, durante las evaluaciones, deberán contar con los servicios de interpretación y los apoyos técnicos necesarios y suficientes”. El Artículo 3 incorpora la importancia de que el desarrollo profesional atienda las demandas individuales de las personas con discapacidad y en el Articulo 29 incluye el inciso e) indicando que cada universidad deberá incluir en sus planes de estudio, en sus actividades de investigación y de extensión la formación y capacitación sobre la problemática de la discapacidad. En este análisis introducimos el aporte de Sebastián Rositto (Rositto, 2012) [4] que analiza la normativa y agrega “Otro avance, es el reconocimiento de los derechos de los estudiantes universitarios en situación de discapacidad a ser evaluados dignamente, ya que se prevé que, durante las evaluaciones, deberán contar con los servicios de interpretación y los apoyos técnicos necesarios y suficientes. Estimamos que quien puede lo más, puede lo menos, por lo tanto, si para las evaluaciones (que son las instancias máximas) cuentan con esos servicios, también debe serlo para las clases en el aula y para cualquier gestión ante una oficina administrativa de dicha institución”. Reconociendo entonces que la legislación vigente establece la obligación del sistema universitario a ser accesible en las dimensiones expresadas al inicio de este apartado (en el espacio educativo, en los recursos didácticos y en los servicios y prácticas de educación). Con respecto a la propuesta de este proyecto de elaborar un modelo de observatorio de accesibilidad para contener información sistematizada de la inserción académica de personas con discapacidad en las aulas, se han analizado la preexistencia de distintos

observatorios en las sedes universitarias. Como antecedentes de Observatorio de accesibilidad encontramos el que existe en la Universidad Nacional de Litoral. También algunas instituciones presentan observatorios para evaluar el acceso a los sistemas web, este es el caso del Observatorio Argentino de Accesibilidad Web [3] que se propone evaluar sitios web particulares usando herramientas de testeo para verificar y validar el nivel de accesibilidad. Para ello utilizan un conjunto de estándares y normas que conforman las Pautas de accesibilidad para entornos web WCAG 2.0. Este observatorio forma parte de una red de observatorios en Iberoamérica

En el orden internacional España es el país que presenta el mayor desarrollo en la materia. INDRA es una empresa multinacional de Tecnología de la Información que impulsa la creación de Cátedras de Tecnología Accesible. Las Cátedras de Tecnologías Accesibles son espacios asociativos entre las universidades y la empresa de tecnología para el desarrollo de actividades de investigación e innovación al servicio de las personas con discapacidad. Los proyectos desarrollados tienen como objetivo utilizar tecnología de punta para dar soluciones innovadoras a las necesidades del colectivo de personas con discapacidad. El Observatorio de accesibilidad tiene por objetivo estudiar y dar a conocer los avances tecnológicos actuales en la materia. Otro antecedente corresponde a la Confederación Española de personas con discapacidad física y orgánica (COCEMFE) que tiene un importante progreso de observatorio de accesibilidad, ya que contempla investigaciones, prácticas y productos tecnológicos para la plena formación de la vida en sociedad de las personas con discapacidad.

En todos los documentos analizados se destaca que la producción de artefactos tecnológicos que permitan la accesibilidad para todos y el desarrollo de nuevos elementos siempre implican un cambio de actitud de los equipos técnicos de ingenieros y diseñadores. Como dice Francisca Negre Bennsar (Negre, 2015) [5] “Debemos ser conscientes de la trascendencia que puede tener para las personas con discapacidad un acceso más sencillo y generalizado a las tecnologías de la información y la comunicación para poder conseguir una igualdad de posibilidades en la educación, el trabajo y la sociedad”.

En el ámbito de la Universidad Tecnológica Nacional se observa en la Facultad Regional Buenos Aires un Área de Orientación y Apoyo para personas con discapacidad (ORADIS) que tiene por propósito coordinar, asesorar, promover y difundir todas aquellas acciones que contribuyan directa o indirectamente a la inclusión de personas con movilidad y/o comunicación reducida, asegurando un equitativo acceso a la educación. Para la formulación de este proyecto se han analizado las estadísticas presentadas por el Servicio Nacional de Rehabilitación dependiente del Ministerio de Salud de la Nación, en el Anuario Estadístico sobre Discapacidad del año 2016. Son datos relevantes para el proyecto los siguientes: el 12,9% de la población tiene discapacidad permanente, en su mayor proporción son mujeres y viven en zonas urbanas, el 59,5 % posee discapacidad visual y el 90% ha asistido a escolarización común.

Líneas de Investigación, Desarrollo e Innovación

Este proyecto se encuentra enmarcado en las líneas de investigación de Tecnología Educativa y dentro de estos en los aspectos de Accesibilidad en los entornos de aprendizaje y en los espacios de formación superior.

Por otro lado, este proyecto incluye el estudio de dimensiones sociológicas y culturales que se asocian a la inserción de las minorías en la educación.

Resultados y Objetivos

El proyecto se encuentra en su fase inicial por lo que los resultados que se presentan son aquellos que se refieren a la investigación de los aspectos educativos a los que los servicios de accesibilidad en la universidad se deben abocar y a la dimensión territorial que se debe abarcar el desarrollo de soluciones tecnológicas apropiadas para la enseñanza de la ingeniería para las personas con discapacidad.

Los objetivos generales de este proyecto son:

1. Desarrollar un sistema de registro de información de personas con discapacidad en estudiantes.
2. Diseñar recursos didácticos que permitan la inclusión de personas con discapacidad en las aulas universitarias de carreras de ingeniería.
3. Generar un observatorio de accesibilidad para la carrera de ingeniería en sistemas de información de la UTN-Facultad Regional Córdoba.

Durante el año pasado se ha participado de las reuniones de formación de la Red IDUC (Red de Inclusión de la discapacidad en las Universidades de Córdoba). Y se ha iniciado la recopilación de información estadística de la participación estudiantil universitaria de estudiantes con discapacidad, en registros propios de la universidad y en bases de datos de acceso público.

En el plano de la carrera de Ingeniería en Sistemas de Información se participó de las reuniones docentes de inicio de año explicando la importancia de la inclusión de la accesibilidad en los trayectos curriculares y en especial en las asignaturas de Trabajo Final.

Se han realizado entrevistas con foco en descubrir espacios no abordados por la enseñanza actual y necesidades de las personas con discapacidad en vista de su experiencia como estudiantes.

Las dimensiones sobre los cuales se iniciará la investigación son:

- Accesibilidad de los entornos web de la universidad.
- Accesibilidad en aulas virtuales.
- Recursos didácticos utilizados en las aulas (exposiciones, proyectores, videos, audios).
- Espacios edilicios dedicados a aulas.
- Propuestas de Incorporación curricular: análisis accesibilidad al momento de exponer requerimientos no funcionales para el desarrollo de software-Norma ISO 25000 (ítem usabilidad).
- Adecuaciones didácticas para no videntes: Reemplazo de Gráficos y tablas por textos planos con el desarrollo de nuevas notaciones de modelado.
- Utilización de letras y símbolos reconocidos por el teclado en fórmulas y expresiones matemáticas y lógicas.
- Análisis de contexto para acompañar a la mejora en la redacción y la ortografía para no videntes.

- Propuestas didácticas para la evaluación de estudiantes no videntes e hipoacústicos.

Formación de Recursos Humanos

Participa como investigador del proyecto la Ing. María Laura Destefanis que se encuentra realizando su tesis de la Maestría en Desarrollo de Entornos Virtuales en la Universidad Nacional de la Patagonia Austral en temática de accesibilidad a los espacios educativos y actualmente se encuentra realizando el curso "Experto en TIC, inclusión y discapacidad. Aplicación práctica de modelos inclusivos con computadoras, tablets y celulares", dictado por Creática FREE Iberoamericana para la Cooperación y con la colaboración de la UNC-FCEFyN.

También participa como investigador tesista la Lic. Elizabeth Jeinson que se encuentra desarrollando su proyecto de tesis de la Maestría en Calidad de Software de la Universidad Nacional de San Luis. Tema: Satisfacción del cliente interno en servicios de tecnología y tecnologías de la información en instituciones educativas de educación superior.

.

Referencias

[1] Peralta Morales, A. (2007). Libro Blanco sobre Universidad y Discapacidad. Real Patronato, Centro de Documentación sobre discapacidad. Madrid.

[2] Ley de Educación Superior, Nro. 24521. República Argentina. Sancionada el 20 de julio de 1995 y Modificación de Ley de Educación Superior. Nro. 25573. República Argentina. Sancionada el 11 de abril de 2002

[3] Observatorio Argentino de Accesibilidad web. www.wcag.com.ar/oia

[4] Rositto, S. (2012) Estado de la educación argentina a tres años de la ratificación de sobre los derechos de las personas con discapacidad. Revista Iberoamericana de educación. Madrid.

[5] Bennar Negre, F. (2015). Internet ¿Un recurso accesible para todos? Universidad de las Islas Baleares.

Observatorio Argentino de Accesibilidad web. www.wcag.com.ar/oia

Desarrollo de Herramientas Informáticas y sus Aplicaciones en el Ámbito Educativo

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.isaia}@est.fi.uncoma.edu.ar, {roberto_laurent}@yahoo.com.ar

Resumen

En este trabajo se presenta una síntesis de los fundamentos y consideraciones para el diseño e implementación de herramientas Informáticas realizadas en el marco del Proyecto de Investigación "Computación Aplicada a las Ciencias y Educación" de la Universidad Nacional del Comahue y algunos avances de las líneas de investigación. Particularmente se presenta la herramienta MCCSOFT y los Objetos de Aprendizaje (OA) desarrollados para la enseñanza y aprendizaje de los Métodos Computacionales. Actualmente estos están siendo utilizados en asignaturas dictadas en las carreras de la Facultad de Informática y Facultad de Economía y Administración, con el objetivo de ayudar al estudiante en la comprensión de los temas del cálculo diferencial y de los métodos numéricos estimulando y reforzando su aprendizaje.

Palabras Claves: Realidad Aumentada, Realidad Virtual, Métodos Computacionales, Objetos de aprendizaje, Cálculo diferencial e Integral, Programación.

Contexto

Las líneas de investigación pertenecen al Proyecto "Computación Aplicada en Ciencias y Educación - F016", Facultad de Informática (FAIF), Universidad Nacional del Comahue (UNCo). Lo integran docentes y estudiantes avanzados de la carrera de Licenciatura en Ciencias de la Computación de la FAIF, docentes de la Facultad de Ingeniería (FI) y docente de la Universidad Católica de Brasilia (UCB) (Brasil).

1. Introducción

Las tecnologías de la información y la comunicación (TIC) han impulsado en las últimas dos décadas el desarrollo de herramientas informáticas a fin de integrarlas a los procesos de enseñanza y aprendizaje en pos de mejorar la calidad educativa en los distintos niveles de educación. En las referencias bibliográficas se describen las dificultades encontradas por los estudiantes en la comprensión y en el aprendizaje de los conceptos teóricos y prácticos del cálculo diferencial e integral en funciones de varias variables. Particularmente, en la visualización de las funciones multivariables y en la

interpretación geométrica de los métodos numéricos [1-3]. Fueron desarrollados en las últimas décadas varios sistemas de cálculo simbólico que pueden ser clasificados como Sistemas de Algebra Computacional (CAS) según sus características, y que en general trabajan con formas algebraicas, numéricas y gráficas. Se puede mencionar que mucho de ellos son software propietarios, requieren licencia para su uso, tienen un alto costo y restricciones en el número de usuarios [4-6]. Si bien existen y se utilizan diferentes software en el dictado de asignaturas de las carreras de la FAIF, se observaron limitaciones en cuanto a la interfaz y a la sintaxis para realizar cálculos en los métodos del cálculo diferencial y en los métodos numéricos. Esto puede deberse a que se necesita tener conocimiento de un gran número de funciones especiales y predefinidas, por lo que su uso para cálculos complejos resulta ser muy tedioso. Además, en la visualización de los resultados por pantalla no se muestran los cálculos intermedios, lo que es muy importante para un mejor entendimiento de los conceptos teóricos y prácticos de los métodos [7-8]. En base a los estudios realizados, y teniendo en cuenta estos problemas se planteó el diseño y desarrollo de una herramienta computacional para suplir estas dificultades y además que cuente con una interfaz de usuario intuitiva y amigable. Para el desarrollo se evaluaron diversos entornos y lenguajes de programación y librerías enfocadas en resolver problemas de matemáticas y computación científica. Se desarrolló la herramienta que fue denominada MCCSOFT como una aplicación Web [9-11]. Fue utilizado el lenguaje Python y para el entorno de desarrollo Web se utilizó el framework Django que implementa la arquitectura Modelo Vista Controlador (MVC) [12].

En la actualidad, los docentes enfrentan en su práctica diaria el desafío de construir el conocimiento con sus estudiantes de la forma más significativa, para lo cual recurren al uso de diversos canales y recursos digitales que le permitan enseñar distintos contenidos y facilitar su comprensión y apropiación. Con el crecimiento desmedido de información en Internet se han generado problemas para ubicar y seleccionar recursos pedagógicos de calidad, ya que, al no tener elementos que describan su contenido o autoría el usuario de se ve expuesto a grandes bancos de información, siendo que muchos de ellos no son de su interés. En base a estos inconvenientes surgieron propuestas de organización de la información bajo el nombre de Objeto de Aprendizaje (OA). Los mismos son materiales de soporte digital y carácter educativo que son diseñados y creados con el propósito de ser reutilizados en distintos ámbitos educativos. En este contexto se sigue principalmente estudiando el potencial de las herramientas de geometría dinámica tridimensional para su desarrollo y que faciliten la práctica docente. El concepto de visualización también adquiere un rol fundamental para la comprensión de los temas del cálculo diferencial e integral en funciones de varias variables. La representación gráfica de una superficie tridimensional generalmente presenta mayor dificultad para su visualización en un ambiente bidimensional como es el lápiz y papel. Esto se debe a que la visualización de funciones de dos variables, a diferencia de lo que ocurre con las funciones en una variable, deben trasladarse del plano al espacio. Es importante que los estudiantes puedan moverse con flexibilidad de una representación a la otra y el uso de recursos tecnológicos contribuye a un mejor logro de este objetivo [13-18].

En base a lo mencionado fueron desarrollados cuatro Objetos de Aprendizaje e implementados utilizando el software de Geometría Dinámica GeoGeogebra [19]. Son ellos: OA1: Curvas de nivel, OA2: Derivadas parciales, OA3: Derivadas Direccionales, OA4: Integrales Dobles.

En la línea “Realidad Aumentada y Realidad Virtual” se sigue avanzando en la investigación con el objetivo de realizar el desarrollo de nuevos recursos para ser

utilizados en diferentes ámbitos, tales como educación, turismo, medicina, patrimonio cultural y entretenimiento. Estas tecnologías emergentes como lo son la Realidad Aumentada (RA) y la Realidad Virtual (RV) utilizan diversos canales y soportes que favorecen la creación de ambientes adecuados para los diferentes estilos de aprendizajes presentes en los niños, favoreciendo una mayor capacidad de percepción, de razonamiento abstracto y espacial. Además se favorece la adquisición y retención del conocimiento en los estudiantes, mediante el uso de diversos medios (audio, animación 3D, etc.) dado que estos permiten interactuar a través de todos los sentidos. Mediante el desarrollo de Proyectos de Extensión se continúa acercando estas tecnologías a Instituciones Educativas de la provincia de Neuquén, donde además de brindar formación para el trabajo con TIC, se propicia el desarrollo colaborativo de nuevos recursos digitales educativos [20-22].

2. Líneas de Investigación y Desarrollo

Las principales líneas que se está trabajando son: Desarrollo y Uso de Recursos TIC, Realidad Aumentada y Realidad Virtual, Métodos Computacionales y Simulación. Las tres líneas propuestas están interrelacionadas, persiguiendo como propósito general acercar la Universidad al medio fomentando el trabajo colaborativo.

3. Resultados Obtenidos/Esperados

De acuerdo a lo descrito anteriormente se desarrolló el software online denominado MCCSOFT, teniendo en cuenta que se respecte la filosofía del software libre. Para el desarrollo se evaluaron diversos entornos y lenguajes de programación, eligiéndose el lenguaje Python por contar con librerías enfocadas en resolver problemas de matemáticas y computación científica. Para el entorno de desarrollo Web se utilizó el framework Django. Se puso énfasis en el diseño de una interfaz de usuario intuitiva y amigable para permitir integrar los contenidos principales del cálculo diferencial e integral con los métodos numéricos. El software fue validado a través de distintos tests entre ellos prueba de usuario a través de encuestas realizadas a los estudiantes, pruebas de desarrollo, de versión, entre otras, y se pudo concluir que MCCSOFT realiza de manera eficiente todas las funciones para las cuales fue diseñado. El mismo permite realizar gráficos en una variable y en 3D. Para funciones en una y varias variables permite calcular derivadas en forma simbólica y numérica, integrales exactas e integrales numéricas. El objetivo principal es su utilización en las asignaturas de las carreras dictadas en la UNCo. Se espera con su utilización lograr mejoras en los procesos de aprendizaje por parte de los estudiantes. Se puede mencionar que actualmente está siendo utilizado en asignaturas del segundo año de las carreras de la FAIF y en materias de la carrera del Profesorado Universitario en Matemática de la Facultad de Economía y Administración. Particularmente en Métodos Computacionales para el Cálculo y Cálculo Numérico. Cabe mencionar que este software fue desarrollado por una alumna integrante del Proyecto de Investigación como parte de su tesis de la carrera Licenciatura en Ciencias de la Computación.

Continuando con los avances de la línea de investigación fueron también desarrollados e implementados cuatro OA por una integrante del grupo de investigación en el marco de una tesis de la carrera de Licenciatura en Ciencias de la Computación. A continuación se menciona en forma sintética los nombres de los OA y su aplicación:

OA1: Curvas de Nivel-Interpretación geométrica
OA2: Derivadas Parciales - Interpretación geométrica
OA3: Derivadas Direccionales - Interpretación geométrica
OA4: Integrales Dobles y Volumen - Aproximación por sumas de Riemann

Se utilizó la metodología CrOA, que propone un proceso guiado a través de una serie de fases o etapa para el diseño y la creación de un OA a partir de preguntas disparadoras que se deben ir contestando y además permiten plasmar el diseño y planificación. Se destaca de otras metodologías por su simplicidad pudiendo ser creada directamente por los docentes [15]. Los contenidos fueron desarrollados a través de la descripción de los objetos geométricos utilizados en el software de Geometría Dinámica GeoGebra [19]. Para los objetos tipo Botón se utilizó la programación de guiones con JavaScript. Los comandos de GeoGebra fueron utilizados para rotar vistas y realizar animaciones, y los objetos de casilla de control para permitir la visualización de distintas figuras. Para la visualización y testeo de los OA se realizaron tests de unidad, módulo, sistema y pruebas de usuario mediante una encuesta a los estudiantes de la asignatura Métodos Computacionales para el Cálculo de la carreras dictadas en la FAIF. El test de aceptación fue del 100% y las funcionalidades consideradas más útiles por los estudiantes fueron: la visualización y animación de los gráficos en 3D, interpretación geométrica y el seguimiento paso a paso de los temas. Se puede concluir que la utilización de los OA ayudaría a una mejor comprensión y relación de los temas teóricos y prácticos. Se propone continuar implementando distintos escenarios utilizando lenguajes de programación Web para la generación de interfaces que permitan la animación y simulación del comportamiento de distintos conceptos matemáticos permitiendo la interacción del usuario a través de la manipulación de diversos parámetros.

Siguiendo con la línea "Desarrollo y Uso de Recursos TIC", se finalizó por parte de una integrante del grupo la tesis de doctorado: "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". El objetivo principal es de generar contribuciones teóricas y metodológicas relacionadas con el uso de torneos de Programación como un recurso didáctico para lograr motivar y mejorar el rendimiento académico de los estudiantes en esta área.

En la línea de Realidad Aumentada y Virtual se continúa investigando la aplicación de estas tecnologías a otros ámbitos, como lo son Museos, Turismo y Medicina, además del educativo en el cual se está trabajando desde el año 2014. Se han desarrollado recursos para el trabajo de contenidos matemáticos, ciencias sociales y naturales, educación física, entre otros. También se han desarrollado recursos digitales provistos de estas tecnologías emergentes en conjunto con organismos tales como el museo de la Universidad Nacional del Comahue y la Subsecretaría de Turismo de la Municipalidad de Neuquén. En el primero caso los objetos 3D desarrollados se expusieron en el marco del evento Misión Dino, realizado en el mes de agosto de 2018. En el segundo caso, se desarrollaron software de Realidad Aumentada y de Realidad Virtual, los cuales se utilizaron en el marco de la fiesta de la Confluencia, llevada a cabo entre los días 8 y 10 de febrero del corriente año. Una de las aplicaciones RA implementadas permite el reconocimiento e interacción con aves que habitan la isla 132 localizada en la ciudad de Neuquén, la misma puede descargarse desde Google Play. Los desarrollos RV contaron además con el trabajo con videos 360º [23-24].

4. Formación de Recursos Humanos

Este proyecto cuenta con once investigadores, entre los cuales se encuentran docentes, estudiantes avanzados de la carrera Licenciatura en Ciencias de la Computación, asesores del país y del exterior. En el año 2018 se finalizó una tesis de doctorado: "Enseñanza de las Ciencias Exactas y Naturales" en la Facultad de Ingeniería (FI), UNCo y también se finalizaron dos tesis de grado de la carrera de Licenciatura en

Ciencias de la Computación: “Desarrollo de una herramienta online basada en software libre como apoyo a la enseñanza de los Métodos Computacionales” y “Diseño e implementación de Objetos de Aprendizaje para la enseñanza de Cálculo multivariable”. Dos docentes investigadores se encuentran realizando su posgrado, habiendo terminado los cursados de las maestrías: “Tecnología Aplicada en Educación – UNLP”, y “Enseñanza de las Ciencias Exactas y Naturales – UNCo)”.

5. Bibliografía

[1] Krishnachandran V.N., Reji C. Joy, K. B. SIPI, More CAS’s in maths classrooms: An urgent imperative, Thechonolgy Enhanced Education (ICTEE), IEEE Internacional Conference on Education, doi: 10.1109/ICTEE.2012.6208645, 2012.

[2] Silva Santana B., Alves da Silva, M. A., Aprendizagem de Cálculo: A partir do uso de software matemáticos, V.1 ed. III CONADU, 2016.

[3] Cavasotto, M, Reflexões sobre as Dificuldades na aprendizagem de cálculo diferencial e integral, III Mostra de Pesquisa da pós-graduação, PUCRS, Brasil, 2010.

[**4**] Mathematica de Wolfram. http://www.wolfram.com/mathematica

[5] MapleSoft. http://www.maplesoft.com

[6]Matlab for Deep Learning. https://www.mathworks.com

[**7]**SoftwareWxmaxima. http://andrejv.github.com/wxmaxima/

[8] Software Scilab. http://www.scilab.org.

[9] C. R., Fundamentals of Web development. Pearson Education, 2015.

[10] Sommerville, I., Software Engineering, 9°ed. Addison-Wesley Publishing Company,USA, 2011

[11] Sharma J., Kumar S. R., Strategies for Web Application Development Methodologies, International Conference on Computing, Communication and Automation (ICCCA), 2016.

[12]Sitio oficial de Django https://www.djangoproject.com

[13] W. R. Fierro and V. A. Bosquez. Design and production of a learning object for university teaching: An experience from theory to practice. ISBN 978-1-5090-6149-5, 2016.

[14] M. A. Mariotti. Introducing students to geometric theorems: how the teacher can exploit the semiotic potential of a dgs. ZDM. The international Journal on Mathematics Education,pp. 441-452, 2013.

[15] C. Sanz, F. Barranquero, and L. Moralejo. Metodología para la creación de objetos de aprendizaje CrOA. http://croa.info.unlp.edu.ar/.

[16] C. Allan, S. Parra, and A. Martins. Objetos de aprendizaje para la interpretación geométrica de métodos numéricos: Uso de geogebra. Revista TE&ET, (20):51_56, ISSN 1850-9959, 2017.

[17] E. Jimenez Lopez, M. Luna Cámara, S. Lopez Cuevas, and R. Peraza Arrollo. Desarrollo de un objeto de aprendizaje para la enseñanza de las matemáticas: el caso de las funciones. Eleventh LACCEI Latin American and Caribbean Conference for Engineering and Technology (LACCEI 2013), 2013.

[18] Lemke, R., Objetos de aprendizagem para o ensino de funções de duas variáveis: um diferencial dinâmico, 2015.

[19] Sitio oficial Software GeoGebra http://www.geogebra.org

[20] Barroso Osuna, J. M., Cabero Almenara, J., García Jiménez, F., Calle Cardoso, F. M., Gallego Pérez, Ó., & Casado Parada, I. Diseño, producción, evaluación y utilización educativa de la realidad aumentada, 2017.

[21] Torres R., 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://revistaseug.ugr.es/index.php/erph/article/view/3395

[22] Fracchia, C., & Martins V.B., A., 2017. Realidad aumentada en la enseñanza primaria: diseño de juegos de mesa para las áreas ciencias sociales y matemáticas. DOCREA, (6), 89-104. https://drive.google.com/file/d/1zlDnTzEP-NsLQbIOv7hV-hX8PG94sIMp/view

[23] http://www.barinoticias.com.ar/index.php?option=com_content&task=view&id=63795&Itemid=2

[24]https://www.copade.gob.ar/contenido.aspx?Id=NOV-8813

Desarrollo de un Dispositivo CANSAT para Fomentar el Acceso a las Tecnologías Espaciales

Leonardo Anchino†, Andrés Torti, Marco Miretti,
Emanuel Bernardi, Gastón Peretti, Rodolfo Podadera
Grupo de Investigación en Robótica y Control (GIROC)
Universidad Tecnológica Nacional
Facultad Regional San Francisco (UTN-FRSFCO)
San Francisco, Córdoba, Argentina
†anchinoleonardo@gmail.com

RESUMEN

En nuestro país, la enseñanza y el acceso a las tecnologías espaciales no está contemplado en los contenidos curriculares de las instituciones educativas de nivel medio. En consecuencia, el acercamiento de los estudiantes a las mismas y su aplicación a través de actividades formativas, es generalmente dificultoso. Esto se debe principalmente a la gran cantidad de conceptos, aprendizajes y equipamiento, que estos sistemas conllevan. Es por ello, que el presente proyecto aborda la investigación y el desarrollo de dispositivos CANSAT para la implementación de una plataforma educativa lista para usar, permitiendo que las instituciones, cualesquiera sea su orientación, se introduzcan progresivamente en la enseñanza de las tecnologías espaciales, sin necesidad de adoptar todos los conocimientos técnicos que esto implica. De esta forma, se posibilita que los estudiantes se enriquezcan con la experiencia obtenida a partir del uso de la misma con un enfoque más general. Promoviendo así, el incentivo y el fortalecimiento de vocaciones científicas.

***Palabras clave**:* educación, electrónica, informática, procesamiento de señales, comunicaciones.

CONTEXTO

Esta propuesta de investigación y desarrollo se enmarca dentro de las áreas Control Automático de Sistemas, Informática y Telecomunicaciones del Departamento de Ingeniería Electrónica de la Universidad Tecnológica Nacional, Facultad Regional San Francisco. Además, el proyecto se corresponde con el PID homologado CCUTNSF0005414, "Desarrollo de una plataforma educativa basada en proyectos CANSAT", y las actividades involucradas en este se llevarán a cabo por los integrantes del Grupo de Investigación en Robótica y Control (GIROC), compuesto de becarios alumnos, docentes investigadores y becarios doctorales.

1. INTRODUCCIÓN

En la actualidad, es común ver como una gran cantidad de jóvenes padecen la incertidumbre de no saber que camino seguir al concluir sus estudios secundarios; la mayoría busca un trabajo, o carreras de corta duración y pocos confían en sus posibilidades de estudiar una carrera de grado. La falta de información, en combinación con la ausencia de interés, constituye una de las principales causas por las que los jóvenes desisten de elegir carreras altamente demandantes, como lo son la Ingeniería Electrónica y la Ingeniería en Sistemas, entre otras. Por otro lado, la necesidad de formar más ingenieros es apremiante y fundamental para resolver desafíos vinculados al desarrollo de la matriz productiva nacional [1, 2].

Entonces, en base a lo previamente expuesto, y con el propósito de fomentar, promover, e

incentivar las vocaciones científicas y tecnológicas en los estudiantes de nivel medio, el presente proyecto pretende facilitar la inclusión de las tecnologías espaciales en la escuela secundaria, como herramienta de promoción de las áreas prioritarias electrónica e informática.

Para ello, se abordará el desarrollo de un pequeño satélite[1] con las dimensiones de una lata de gaseosa, también llamado CANSAT [4, 5]. El dispositivo consiste en un satélite artificial, no orbital, que es lanzado a grandes alturas para la recolección de datos durante su ascenso, y posterior descenso controlado por un paracaídas. Dicha información se transmite en tiempo real hacia una estación terrena, para su posterior análisis y procesamiento. Esto es, para abordar el desarrollo de la plataforma CANSAT se evaluarán las distintas alternativas tecnológicas disponibles [7, 8, 9], y al mismo tiempo se tendrá en cuenta el reglamento vigente en el ámbito internacional para este tipo de competencias [3].

Se entiende como plataforma CANSAT al conjunto de software y hardware, comprendido por el satélite, su estación terrena y su programa de computadora respectivo, como medio diseñado especialmente para permitir a los alumnos del nivel medio insertarse en el mundo de las tecnologías espaciales a través de la programación, la electrónica, el uso de estadísticas, la interpretación de datos, etc.

Por último, es importante destacar que este proyecto surge en base a la necesidad expuesta entre docentes y directivos de la UTN Facultad Regional San Francisco y sus pares de la escuela secundaria IPET N° 50 Emilio F. Olmos, también de la ciudad de San Francisco. Donde la primera proveerá del asesoramiento técnico y diseño de la plataforma CANSAT y la segunda será la encargada de generar acciones formativas en torno a dicha plataforma, para en el futuro inmediato conformar competencias entre los estudiantes que favorezcan aún más el acercamiento de los jóvenes a la ciencia y tecnología, específicamente aquellas que son afín a las tecnologías espaciales.

La plataforma CANSAT se compone de tres partes fundamentales:

1. CANSAT: pequeño satélite no orbital del tamaño de una lata de gaseosa.
2. Estación Terrena: módulo terrestre encargado de establecer el enlace de comunicación con el satélite.
3. Programa de control: permite configurar, administrar y monitorizar el satélite en todo momento.

En la industria satelital, cada satélite puesto en órbita, comercial o no, es proyectado y diseñado en torno a una misión específica. Habitualmente, estas tienen una duración de varios años, y dependen de las características del satélite y del objetivo específico de la misión. Además, los satélites cuentan con la capacidad de intercambiar datos de telemetría, e información de la misión, con estaciones terrenas de control. Entre otras capacidades, los satélites científicos, a menudo, están equipados con diversos sensores para el estudio de fenómenos atmosféricos, la medición de niveles de contaminación, la concentración de gases en la atmósfera, la monitorización de catástrofes naturales, el análisis de cultivos, etc.

Entonces, las misiones de la plataforma CANSAT son similares a la de un satélite orbital solo que, en este caso, el mismo es no orbital y las misiones duran unos minutos, en lugar de varios años. Para efectuar una misión CANSAT, se lo eleva a una altura aproximada de 1000 m, seguido por un descenso controlado donde, en tiempo real, se realiza la recolección de los datos suministrados por los sensores a bordo del satélite. Dicha información se recibe a través de la estación terrena y se visualiza en un programa de computadora, desde el cual se los exporta

[1] *Si bien se los denomina «satélites», no lo son en el sentido estricto de su definición como cuerpo que gira alrededor de un planeta.*

para un análisis en detalle de los mismos. Las misiones realizadas con la plataforma CANSAT, entonces, tienen la misma metodología de planificación que las realizadas en satélites orbitales, permitiendo así, que los estudiantes se acerquen a la experiencia de planificar y desarrollar misiones, para luego proceder con el análisis de los resultados recogidos en vuelo, de manera simple y estructurada. De este modo, es posible involucrar a múltiples asignaturas escolares, tales como matemática, física y ciencias naturales, en el planeamiento de las misiones y en el posterior estudio de los datos obtenidos. Como consecuencia, se enriquecen las investigaciones y experiencias prácticas en complemento de los conceptos teóricos, enseñados en el aula, fomentando el trabajo en equipo, la redacción de informes y el diseño de proyectos basados en las experiencias realizadas.

Las medidas del satélite CANSAT están estandarizadas y para el desarrollo del mismo se tomó como base el reglamento [3] que utiliza la Agencia Espacial Europea en sus competencias anuales de plataformas CANSAT. Se optó por este reglamento para permitir que la plataforma sea factible de utilización, sin modificación alguna, en competencias de este tipo. Posibilitando la creación de este tipo de competencias a nivel nacional e incrementando el incentivo a la participación de estudiantes.

El objetivo principal del proyecto es el de proveer un satélite tipo CANSAT completamente funcional y listo para ser empleado en conjunto con una estación terrena de comunicaciones. Además, dicha plataforma educativa permite, no solo su utilización sin modificación alguna, sino que posee la capacidad de ser expandida fácilmente para agregar o mejorar las funcionalidades básicas, dando lugar a la experimentación y al desarrollo de la creatividad en los estudiantes.

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 satélite y la estación terrena.
- Evaluación de las tecnologías a utilizar en el desarrollo del software multiplataforma para computadora.
- Desarrollo del software y hardware necesarios para el satélite y la estación terrena.
- Generación de documentación adecuada para la discusión interna de resultados.

3. OBJETIVOS

Objetivo General:

Desarrollar una plataforma lista para usar, fácil de replicar, de código y diseño libre y abierto, con capacidad de expandirse a través de módulos de simple conexión y programación que permita, con su funcionalidad base, obtener un CANSAT completamente funcional.

Objetivos específicos:

- Seguir la normativa impuesta por la Agencia Espacial Europea (ESA).
- Diseñar un sistema base de gestión de baterías, unidad de procesamiento, módulos sensores (GPS, acelerómetro, magnetómetro, giroscopio, sensor de temperatura, humedad y presión) e interfaz de comunicación.
- Diseñar un método de conexión que permita añadir fácilmente módulos de expansión internos que amplíen las capacidades del CANSAT.
- Diseñar protecciones para evitar el daño del módulo principal al conectar módulos de expansión defectuosos o incorrectamente conectados.
- Diseñar módulos de expansión para el CANSAT que se conecten al sistema principal y sirvan de ejemplo para los estudiantes.

- Diseñar una estación terrena capaz de comunicarse con el CANSAT, permitiendo la recolección de los datos generados por los sensores.
- Construir las placas de circuito impreso correspondientes.
- Diseñar un software de administración y visualización, en computadora, que se conecte la estación terrena y permita observar los datos recibidos en tiempo real, su almacenamiento, exportarlos para su posterior análisis.
- Crear librerías que simplifiquen el uso de la plataforma.

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 un satélite tipo CanSat acorde a las especificaciones de la Agencia Espacial Europea.
- Desarrollo de una estación terrena con su software para computadora correspondiente capaz de recibir y procesar los datos provenientes del satélite.
- Desarrollo de librerías de software para facilitar el uso de la plataforma.
- 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.
- Generación de documentación para la discusión y exposición de resultados.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo abocado al proyecto se encuentra integrado al Grupo de Investigación en Robótica y Control (GIROC), el cual está conformado por su director el Mg. Ing. Rodolfo Podadera, los investigadores de apoyo Mg. Ing. Gastón Peretti, Ing. Emanuel Bernardi y los alumnos avanzados Marco Miretti, Andrés Torti y Leonardo Anchino. 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] UTN-BA, “Esfuerzos en todo el país para promover el ingreso a la UTN”. URL: https://www.frba.utn.edu.ar/esfuerzos-en-todo-el-pais-para-promover-el-ingreso-a-la-utn/ (visitado 08-03-2019)

[2] ITBA, “Apostar a las ingenierías, las carreras del futuro”. URL: https://www.itba.edu.ar/la-universidad/prensa/apostar-a-las-ingenierias-las-carreras-del-futuro/ (visitado 08-03-2019)

[3] European Space Agency, (2019) “CANSAT, 2018-2019 Guidelines”.

[4] N. Sako, Y. Tsuda, S. Ota, T. Eishima, T. Yamamoto, I. Ikeda, S. Nakasuka, (2001). “Cansat suborbital launch experiment-university educational space program using can sized pico-satellite”. *acta astronautica*, 48 (5-12), 767-776.

[5] S. Bulut, M. Gül, C. Beker, I. İpek, O. Koçulu, C Topaloğlu, C. Tüfekçi, (2013). “Model satellite design for CanSat Competition”. In *2013 6th International Conference on Recent Advances in Space Technologies (RAST)* (pp. 913-917). IEEE.

[6] A. Nylund, J. Antonsen, (2007). “CanSat-general introduction and educational advantages”. In *Proceedings of the 18th ESA Symposium on European Rocket and Balloon Programmes and Related Research, Visby*.

[7] Richard Barry (2016). Mastering the FreeRTOS Real Time Kernel, A Hands On Tutorial Guide.

[8] Amazon Web Services (2017). The FreeRTOS Reference Manual.

[9] Horowitz, P., Hill, W., & Robinson, I. (1980). The art of electronics (Vol. 2). Cambridge: Cambridge university press.

"Desarrollo, aplicación y evaluación de herramientas didácticas con RA en el sistema educativo de la provincia de Jujuy"

María del Pilar Gálvez[1, 2], Nélida R. Cáceres[1], Ana C. Tolaba[1], Laura Villarrubia[1], Eduardo Graneros[2], José Quispe[1], Fabiana Chiquello[2], Mariela Calisaya[2], Penelope Camara[2]
[1]Facultad de Ingeniería, Universidad Nacional de Jujuy
Italo Palanca 20 San Salvador de Jujuy – 0388 4221576
[2]Universidad Católica de Santiago del Estero – Departamento Académico San Salvador
Lavalle 333 San Salvador de Jujuy – 0388 4236139
mpgalvezd@yahoo.com.ar

RESUMEN

Las computadoras y las nuevas tecnologías en la educación pueden ser vistas como herramientas y materiales de construcción que facilitan el aprendizaje y el desarrollo de habilidades, constituyendo distintas formas de aprender para los estudiantes y no deben ser consideradas como un fin, sino como un medio.

La Realidad Aumentada (RA) es una tecnología novedosa en el ámbito educativo de la provincia de Jujuy y puede ser utilizada para promover, facilitar y enriquecer la comprensión de temas y conceptos propios e importantes de las diferentes asignaturas curriculares dentro del proceso educativo. En este proyecto se propone generar una instancia de trabajo e intercambio entre el equipo de investigación y actores institucionales del sistema educativo provincial para el desarrollo, aplicación y evaluación de herramientas didácticas con RA.

Palabras clave: *TIC's, Realidad Aumentada, Educación.*

CONTEXTO

La presente investigación se enmarca en el proyecto "Realidad Aumentada: innovación educativa en las aulas jujeñas", categoría A, aprobado por la Secretaría de Ciencia y Técnica de la Universidad Nacional de Jujuy (UNJu) y se encuentra bajo incentivo. Código de proyecto 08D/144.

El proyecto es llevado a cabo por GIDIS (Grupo de Investigación y Desarrollo en Ingeniería de Software) de la Facultad de Ingeniería de la UNJu.

Participa además en el desarrollo de este proyecto la Universidad Católica de Santiago del Estero, Departamento Académico San Salvador. Con la incorporación de docentes investigadores de la mencionada casa de estudios, se constituyó un equipo de trabajo más amplio, diversificado, interdisciplinario beneficiando al proyecto con el aporte del área de conocimiento de todos sus integrantes.

1. INTRODUCCIÓN

Las instituciones educativas y sus prácticas han quedado atrapadas en el tiempo, en el sentido de su oferta cultural, de aprendizaje y de enseñanza frente a nuevo sujetos producto de las condiciones sociales, culturales y fundamentalmente tecnológicas, como son los estudiantes que corresponden a la denominada generación NETs [1]. Este autor indica, que estos estudiantes, al estar expuestos a mucha información en las redes cualquier situación de enseñanza puede resultarles aburrida y continuamente buscan ámbitos y experiencias de aprendizaje divertidas. Requieren de instrucciones precisas, de planes detallados de trabajo que impliquen recompensas inmediatas, en consecuencia la responsabilidad y compromiso surgen cuando encuentran sentido en lo que hacen, cuando son útiles para alguien o para algo.

Coincidimos con Larry Cuban [2] en reconocer que "sin poner atención a las condiciones del lugar de trabajo de los docentes y sin el reconocimiento de los saberes que traen los alumnos al aula, hay poca esperanza de que las nuevas tecnologías

tengan más que un mínimo impacto en la enseñanza y el aprendizaje".

Frente a este escenario, la oferta educativa debería abandonar su marca tradicional y apuntar a la generación de saberes y formas de enseñarlos que tomen en cuenta las características de estos nuevos sujetos que habitan las aulas. Los materiales didácticos que operan con el lenguaje multimedial se acercan más a la experiencia de estos estudiantes, lo que conlleva un aumento en la capacidad de retención de la información y una mejora en los resultados pedagógicos.

Las computadoras y las nuevas tecnologías en la educación pueden ser vistas como herramientas y materiales de construcción que facilitan el aprendizaje y el desarrollo de habilidades; constituyendo distintas formas de aprender de los alumnos y no deben ser consideradas como un fin, sino como un medio. Se busca utilizar estas herramientas para promover, facilitar y enriquecer la comprensión de temas y conceptos propios e importantes de las diferentes asignaturas curriculares dentro del proceso educativo.

En este sentido Realidad Aumentada (RA) como nueva tecnología puede incorporarse en el proceso de enseñanza y aprendizaje dentro del sistema Educativo, mediante la combinación de diferentes tipos de materiales (auditivos, visuales y audiovisuales).

La Realidad Aumentada propone superponer o añadir, en tiempo real, elementos virtuales al campo de visión de una persona, de modo tal que la información requerida esté presente de una manera natural para el usuario. Estos elementos consisten en imágenes y textos visuales generados por computadora, llevando a que el usuario no se sumerja completamente en un mundo virtual sino en una mezcla de éste con el mundo real, manteniéndolo en contacto con la realidad y en ocasiones permitiéndole interactuar con objetos virtuales. Esto es diferente de la Realidad Virtual en donde se generar un mundo artificial el que reemplaza por completo al mundo real, teniendo por objetivo principal que el usuario pueda interactuar y sentirse inmerso dentro de éste mundo generado [3].

Para hacer uso de esta tecnología se debe contar mínimamente con: (Fig. 1)

- *Cámara Web:* necesaria para la captura de las escenas del mundo real.
- *Software de Realidad Aumentada:* la cámara web genera un video del entorno existente en tiempo real enviándolo al software de RA, quien es el encargado de procesar la información adquirida por la cámara, los datos almacenados en su base de datos, la información proporcionada por los marcadores (necesarios para indicar donde mostrar la información virtual generada) así como la posición, orientación y marcas de reconocimiento. Finalmente, sobre la imagen capturada y mediante librerías externas, serán dibujados los objetos 3D de modo que aparezcan en la posición, orientación y tamaño correspondiente al punto de vista de la cámara [4].
- *Dispositivo de Visualización:* a través de éste, pantalla o monitor, se proyectará la suma de lo real y lo virtual, lo cual consiste en añadir información de forma numérica, textual o bien objetos que no estaban presentes en la escena original, conformando la RA. Esta composición es la que el usuario final verá reflejada en el dispositivo de salida.

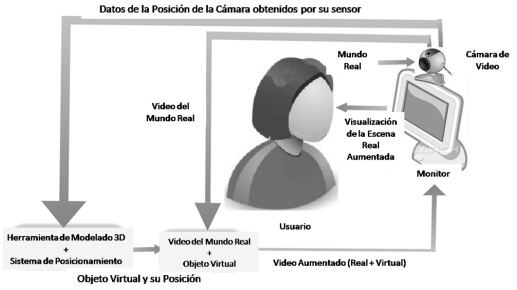


Figura 1. Sistema de Visión por Video basado en Monitor.

2. LÍNEAS DE INVESTIGACIÓN y DESARROLLO

- Proponer a los educadores de distintos niveles de enseñanza de la Provincia de Jujuy nuevas propuestas de TIC´s para mejorar el proceso de enseñanza y

aprendizaje con la introducción de la Realidad Aumentada.

- Proponer las competencias necesarias para el trabajo y desarrollo de aplicaciones usando RA.
- Desarrollar aplicaciones de RA como recursos metodológicos novedosos para facilitar los procesos de enseñanza y aprendizaje.
- Proponer desarrollos para RA en el contexto de la Ingeniería de Software.

3. RESULTADOS OBTENIDOS

Resultados obtenidos en relación a las líneas de investigación y trabajo del proyecto:

- Se llevo a cabo el dictado de un curso de capacitación para docentes de escuelas primarias de Jujuy con el auspicio de Secretaria de Ciencia y Técnica del Ministerio de Educación de la provincia, Resolución Ministerial 10412-E-18. El propósito de la capacitación fue brindar los conocimientos básicos de RA, capacitar a los docentes en el uso de aplicaciones educativas de RA, contextualizar la enseñanza de los estudiantes NET´s y diseñar y llevar a cabo propuestas de experiencias prácticas de uso de aplicaciones de RA en las diferentes áreas disciplinares curriculares en las cuales se desempeñan con la colaboración y soporte de GIDIS.
- Se realizó un taller de capacitación en creación de aplicaciones de RA para educación utilizando Aurasma (HP-REVEAL) [5].

Respondieron a la convocatoria 53 docentes de escuelas primarias públicas y privadas de Jujuy de los cuales 27 realizaron el diseño y llevaron a cabo una experiencia práctica en aula con sus estudiantes usando aplicaciones de RA.

Las conclusiones de los docentes que llevaron a cabo experiencias prácticas ponen de manifiesto una mejora en la comprensión de los temas impartidos, facilidad para aprender de forma casi natural los temas de estudio y que los estudiantes se mostraron motivados, participativos y colaborativos con sus compañeros en el trabajo solicitado por el docente.

No hubo problemas con el uso de los dispositivos (computadoras, celulares y tablet´s) necesarios para el trabajo, sin considerar la edad de los estudiantes conocían como usarlos y se ayudaron entre ellos para lograr la tarea propuesta sin requerir la atención del docente [6].

- Se propusieron las competencias necesarias para trabajar con RA tanto para el uso de aplicaciones en el contexto educativo como para desarrollar aplicaciones educativas [7].
- Se desarrollaron aplicaciones de RA educativas para dispositivos móviles utilizándose software de tipo libre:
  - Geografía interactiva: permite aprender a identificar los departamentos que componen la provincia de Jujuy, su capital, región a la que pertenece y cantidad de habitantes que posee [8].
  - Provincias y capitales de Argentina con Realidad Aumentada [9].
  - Anatomía con RA para el estudio de los sistemas óseo y circulatorio del cuerpo humano [10].
- En el ámbito de la Ingeniería de Software se desarrolló una metodología que permite diseñar e implementar aplicaciones educativas con Realidad Aumentada en donde se combinan el trabajo ágil, la reutilización, la calidad y la interacción activa de docentes con el equipo de desarrollo [11].

4. FORMACIÓN DE RECURSOS HUMANOS

Estructura del equipo de trabajo:

Directora: Mg. María del Pilar Gálvez. Categoría de Investigación III.

Asesora en educación: Dra. Prof en Ciencias de la Educación Jorgelina Argañaraz.

Investigadores:

- Ing. Nélida R. Cáceres. Categoría de Investigación IV.
- Ing. Ana Carolina Tolaba. Categoría de Investigación V.

- Ing. Laura Villarrubia. Categoría de Investigación IV.
- Ing. Eduardo Graneros.
- Ing. José Quispe.
- Ing. Fabiana Chiquello.
- Ing. Mariela Calisaya.

Estudiante de la carrera de Ing. Informática:

- Penélope Camara.

Becarios Cin:

- Daniel Lamas: "*Batalla Naval Online con Realidad Aumentada"*. Desarrollo de video juego con RA [12]. Segundo lugar en trabajos de estudiantes investigadores, CONAIISI 2017.
- Eduardo Hinojosa: "Aplicación de Realidad Aumentada para educación: 23 de agosto (Éxodo Jujeño)". En este desarrollo se combina el despliegue de distintos tipos de información: texto, imágenes en 3D, mapas, filmaciones realizadas con Dron y filmaciones de tipo documentales.

Trabajos Finales de Grado:

- Adrian Flores: "Desarrollo de prototipo de identificación de una especie animal por medio de sus sonidos característicos". Trabajo Final de grado para la obtención del título de Ing. en Informática.

5. BIBLIOGRAFIA

[1] Córica, J. L., Dinerstein, P. *Diseño Curricular y Nuevas Generaciones: Incorporando a la Generación Net*. Editorial Virtual Argentina. Mendoza, 2009.

[2] Cuban, L. *The Myth of Failed School Reform*. Disponible en: http://larrycuban.wordpress.com/2011/02/02/the-myth-of-failed-school-reform-part-1/ y http://larrycuban.wordpress.com /2011/02/04/myth-of-failed-school-reform-part-2/.

[3] Abásolo, M. J., Rodríguez, R. A. *Aplicaciones de la Realidad Aumentada: Enfoques Posibles para Entornos Educativos*. La Plata, 2008.

[4] Kato, H., Bilinghurst, M. *Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System*. 2do Workshop Internacional sobre Realidad Aumentada. San Francisco, 1999.

[5] "Hewlett Packard Reveal", https://studio.hpreveal.com. Acceso Septiembre, 2018.

[6] Camacho Montaño, C. R., Gálvez Díaz, M. del P. *Prácticas de la Enseñanza y la Realidad Aumentada: Aportes Universitarios para la Construcción de Recursos Didácticos Innovadores*. Seminario Internacional CRISCOS: Por la Integración Regional: de los seminarios al Consejo de Rectores. Jujuy, 2018.

[7] Gálvez, M. del P., Camacho, R. *Desarrollo de Competencias en el Aula para el Abordaje de la Realidad Aumentada: Extensión Universitaria para la Capacitación en Servicio*. XII Jornadas de Ciencia y Tecnología de Facultades de Ingeniería del NOA. Catamarca, 2017.

[8] Quispe, J., Gálvez, P. *Realidad Aumentada aplicada al aprendizaje de Geografía en escuelas primarias.* VI Seminario Argentina – Brasil de Tecnologías de la Información y la comunicación. Concepción del Uruguay, 2018.

[9] Quispe, J. *Provincias y capitales de Argentina con Realidad Aumentada*. 6to Congreso Nacional de Ingeniería Informática /Sistemas de Información. Mar del Plata, 2018.

[10] Veramedi, B., Cámara, A. & Calisaya, M. *Realidad Aumentada Innovación en la Enseñanza del Cuerpo Humano en la Educación Primaria de Jujuy*. IV Jornadas de Jóvenes Investigadores. San Salvador de Jujuy, 2018.

[11] Cáceres, N., Tolaba, A. C. *Metodología de Desarrollo para Aplicaciones con Realidad Aumentada en la Educación.* 5to Congreso Nacional de Ingeniería Informática /Sistemas de Información. Santa Fe, Argentina, 2017.

[12] Lamas, D. A. *Batalla Naval Online con Realidad Aumentada.* 5to Congreso Nacional de Ingeniería Informática /Sistemas de Información. Santa Fe, Argentina, 2017.

Detección de Problemas de Aprendizaje y Recomendación de Tutores para Entornos U-Learning

Juárez Luciano Gastón, Beatriz Fernández-Reuter, Elena Durán
Instituto de Investigaciones en Informática y Sistemas de Información (IIISI)
Facultad de Ciencias Exactas y Tecnologías (FCEyT)
Universidad Nacional de Santiago del Estero (UNSE), Santiago del Estero
lucianojuarezgaston@gmail.com; bfreuter@unse.edu.ar; eduran@unse.edu.ar

RESUMEN

En los últimos años ha surgido un nuevo escenario educativo caracterizado por un conjunto de actividades formativas accesibles en cualquier lugar y desde cualquier dispositivo, el aprendizaje ubicuo (u-learning). Sin embargo, cuando un estudiante presenta algún problema en su aprendizaje, muchas veces no sabe a quién recurrir o desconoce que la ayuda está más cerca de lo que cree. Es por esto que se ha iniciado una línea de investigación enfocada en los sistemas de recomendación para entornos de aprendizaje ubicuo. En particular esta investigación se propone diseñar un sistema multiagente que tenga en cuenta los datos del contexto del estudiante, como la ubicación y el tiempo disponible, para recomendarle un tutor adecuado.

CONTEXTO

Esta línea de investigación se desarrolla en el marco del proyecto de investigación "Métodos y técnicas para desarrollo de aplicaciones ubicuas", código 23/C139, correspondiente a la convocatoria 2017-2020 de la Secretaría de Ciencia y Tecnología de la UNSE (SICYT – UNSE). El mencionado proyecto tiene entre sus objetivos "*Construir y evaluar el modelo de contexto ubicuo", "Evaluar y proponer técnicas y métodos para el razonamiento y la adaptación de aplicaciones ubicuas" y "Aplicar las técnicas, métodos y estrategias propuestos en la construcción de aplicaciones ubicuas".* Con el propósito de dar cumplimiento a estos objetivos específicos, se identificaron y analizaron problemáticas locales y regionales que potencialmente pueden demandar la construcción de aplicaciones ubicuas. Entre esas problemáticas se identificó el apoyo on-line para reconocer tutores de aprendizaje a quién el estudiante pueda realizar una consulta.

1. INTRODUCCIÓN

En los últimos años ha surgido un nuevo escenario educativo caracterizado por un conjunto de actividades formativas accesibles en cualquier lugar y desde cualquier dispositivo, el aprendizaje ubicuo (u-learning). El uso de dispositivos como, teléfonos móviles y tabletas, crean nuevas oportunidades para los estudiantes, quienes están permanentemente conectados. De esta manera, se puede acceder a contenido educativo, y las interacciones se pueden llevar a cabo, donde los estudiantes lo necesitan sin restricciones de espacio o tiempo (Graf and Kinshuk 2008). Sin embargo, cuando un estudiante presenta algún problema en su aprendizaje, muchas veces no sabe a quién recurrir, o desconoce que la ayuda está más cerca de lo que cree. En este sentido, un sistema de recomendación, que tenga en cuenta datos del contexto del estudiante, como la ubicación y el tiempo disponible, lo puede asesorar para encontrar un tutor adecuado. Existen numerosos trabajos que proponen sistemas de recomendación para entornos de

aprendizaje ubicuo. Entre los que se pueden mencionar a (Boudabous et al. 2017) quienes proponen un sistema multiagente que proporciona soporte a los estudiantes, mediante el uso de dispositivos móviles, en un sistema de aprendizaje ubicuo implementado dentro de un hospital. Este trabajo se centra en proveer contenido educativo personalizado a demanda. En (Salazar et al. 2015) presentan un sistema multi-agente para u-learning sensible al contexto que ofrece planificación adaptativa de cursos virtuales, evaluación personalizada de cursos, selección de objetos de aprendizaje según el perfil del estudiante, búsqueda de objetos de aprendizaje en repositorios, búsqueda de asistentes de aprendizaje por temas y el acceso a diferentes actividades aprendizaje colaborativo sensibles al contexto. (Ayoola and Mangina 2014) presentan PULP, una plataforma de aprendizaje ubicua personalizada diseñada para proporcionar contenido personalizado, actividades y servicios de colaboración, accesibles en cualquier momento y lugar. En (Tan et al. 2009) se presenta la arquitectura de un sistema multiagente con varios objetivos de adaptación. Por un lado, el sistema permite a los estudiantes crear contenido de aprendizaje basado en la ubicación y compartirlo con otros compañeros. También permite crear grupos de aprendizaje dinámicamente al identificar las ubicaciones geográficas de los dispositivos móviles de los estudiantes y otros factores de aprendizaje, como el historial de aprendizaje, el rendimiento previo, el nivel educativo, el estilo de aprendizaje y los intereses. Además, todo el contenido presentado al alumno se adapta según las características de su dispositivo móvil. En (Duran and Álvarez 2017) se ha presentado un método para recomendar a los estudiantes, expertos en el tema que quieren aprender, considerando para ello las experiencias de esos expertos con otros estudiantes, la disponibilidad de estos y la cercanía física con el estudiante para el caso de una consulta presencial. En el marco de esta investigación se propone integrar dicho método con una arquitectura multiagente, que permita monitorear las tareas del estudiante con el fin de detectar problemas de aprendizaje y recomendar las personas que pueden ayudarlo. Un sistema multi-agente es un consorcio de múltiples agentes trabajando juntos, donde cada agente se especializa en una función determinada (Sajja and Akerkar 2012). Estos agentes son capaces de actuar de manera precisa con el fin de realizar tareas en representación de su usuario (Nwana 1996). Para esto, se sitúan en un ambiente determinado y son capaces de ejecutar flexiblemente acciones autónomas con el fin de cumplir sus objetivos (Wooldridge 2002).

2. LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Para la detección de problemas de aprendizaje y recomendación de tutores en entornos de u-learning se deben abordar las siguientes cuestiones:

- Definición de mecanismos que permita detectar problemas de aprendizaje en entornos de aprendizaje ubicuo. Para esto se evaluarán los técnicas y métodos existentes, determinando los más adecuados para su aplicación en entornos de aprendizaje ubicuo.
- Establecimiento de los datos contextuales que se tendrán en cuenta para la búsqueda de tutores a recomendar. Se analizarán diferentes sistemas de recomendación aplicados a entornos de u-learning, para determinar los datos contextuales que se deben tener en cuenta para la recomendación de tutores.
- Integración del método de recomendación propuesto por (Duran and Álvarez 2017) con los sistemas multiagente para combinar la detección de problemas de aprendizaje y la

recomendación de tutores en entornos de aprendizaje ubicuo.

3. OBJETIVOS Y RESULTADOS ESPERADO

El objetivo principal de esta línea de investigación es la de diseñar un sistema multiagente para la detección de problemas de aprendizaje y la recomendación personalizada de tutores en entornos de aprendizaje ubicuo.

Los objetivos específicos de esta investigación son:

- Analizar y evaluar diferentes modelos multiagente para la detección de problemas de aprendizaje y recomendación de tutores en entornos de aprendizaje ubicuo.
- Diseñar y construir los agentes para la detección de problemas de aprendizaje y recomendación de tutores.
- Diseñar y construir el prototipo de sistema que brinde recomendaciones de tutores en entornos de aprendizaje ubicuo.

Los resultados que se esperan obtener de esta investigación, conforme a los objetivos, son:

- Estado del arte referido a los diferentes modelos multiagente para la detección de problemas de aprendizaje y recomendación de tutores en entornos de aprendizaje ubicuo.
- Agentes inteligentes para la detección de problemas de aprendizaje y recomendación de tutores.
- Prototipo de sistema de recomendación de tutores en entornos de aprendizaje ubicuo.

4. FORMACIÓN DE RECURSOS HUMANOS

En el marco de esta línea de investigación se está formando un Becario de Grado con Beca de Investigación para Estudiantes, otorgada por la Consejo de Investigación Científica y Tecnológica de la Universidad Nacional de Santiago del Estero (CICyT-UNSE), período 2019, y se está desarrollando un trabajo final de graduación de la carrera Licenciatura en Sistemas de Información de la Universidad Nacional de Santiago del Estero

5. BIBLIOGRÁFIA

AYOOLA, O.L. AND MANGINA, E. 2014. Personalisation of a U-Learning Environment for Third Level Education. Yuksekogretim Dergisi 4, 1, 54–60.

BOUDABOUS, S., KAZAR, O., AND LAOUAR, M.R. 2017. Towards a hospital ubiquitous learning system based agents. Proceedings of the International Conference on Computing for Engineering and Sciences, ICCES ’17, 56–62

DURAN, E.B. AND ALVAREZ, M.M. 2017. Method for generating expert recommendations to advise students on ubiquitous learning experiences. Proceedings - International Conference of the Chilean Computer Science Society, SCCC, 1–8.

GRAF, S. AND KINSHUK. 2008. Adaptivity and Personalization in Ubiquitous Learning Systems. 4th Symposium of the Workgroup Human-Computer Interaction and Usability Engineering - USAB 2008, Springer 2008, 331–338.

NWANA, H. 1996. Software Agents: An Overview. Knowledge Engineering Review 11, 3, 205–244.

SAJJA, P.S. AND AKERKAR, R. 2012. Intelligent Technologies for Web Applications. Chapman & Hall/CRC.

SALAZAR, O.M., OVALLE, D.A., AND DUQUE, N.D. 2015. Adaptive and personalized educational ubiquitous multi-agent system using context-awareness services and mobile devices. In: P. Zaphiris and A. Ioannou, eds., Learning and Collaboration Technologies. Springer International Publishing, 301–312.

TAN, Q., KINSHUK, KUO, Y., ET AL. 2009. Location-Based Adaptive Mobile Learning Research Framework and Topics. 2009 International Conference on Computational Science and Engineering, 140–147.

WOOLDRIDGE, M. 2002. Introduction to MultiAgent Systems. Wiley & Sons.

WOOLDRIDGE, M. AND JENNINGS, N.R. 1995. Intelligent Agents: Theory and Practice. Knowledge Engineering Review 10, 115–152.

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, Mamani Jonatan, Aguilar Santiago

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, jonush88@gmail.com, santeex@gmail.com

RESUMEN

En la actual sociedad del conocimiento, la incorporación de las nuevas tecnologías de la información y de la comunicación en el ámbito educativo tienen un impacto significativo en el desarrollo del aprendizaje de los estudiantes y en el fortalecimiento de sus competencias, beneficiando su posterior inserción en el sector productivo [1].

La inclusión de áreas de conocimiento como programación y robótica resultan elementales en la formación de profesionales del futuro, con independencia de la disciplina elegida.

No obstante, esta incorporación presenta importantes desafíos al sistema educativo.

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 utilizadas previamente 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.

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 30/04/2019.

INTRODUCCIÓN

La enseñanza de Informática es actualmente una preocupación global; países como Nueva Zelanda, Estonia,

Japón, Finlandia y Reino Unido, han actualizado sus currículos escolares incluyendo la enseñanza de la programación en las escuelas [3][4]. Algunos 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 codeacademy [5].

En Argentina, a lo largo de las últimas décadas, el uso de TICs se ha ido incorporando a las prácticas educativas en los distintos niveles de la escolaridad obligatoria [6].

Incorporar la disciplina informática en los diseños escolares, desde niveles iniciales, 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 capacidad de abstracción imprescindible para manejar la complejidad, a la generalización mediante el descubrimiento de patrones y similitudes y a la evaluación [7].

Estas actividades, que identifican al "pensamiento computacional", son ampliamente aplicadas en la resolución de problemas usando computadoras y en la comprensión de los sistemas más allá del espacio escolar. A su vez, la enseñanza y el aprendizaje de "programación" es un aspecto de la Informática que en la escuela se percibe como el más desafiante [8] [9].

Por otra parte, 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. Esto se evidencia, al menos en las universidades argentinas, en elevados índices de fracaso en las materias iniciales que tratan estos temas.

La importancia de trabajar sobre la reducción de la deserción y un eventual aumento de matrícula es aún más crítico en la disciplina, dada la escasez de profesionales, tanto en la provincia como en el país. Este hecho se percibe a diario en nuestra sociedad y está avalado por acciones establecidas desde distintos entes nacionales en los últimos años, como por ejemplo el Programa Robótica y Tecnología para Educar, del Ministerio de Ciencia, Tecnología e Innovación Productiva de la Nación, el Plan Nacional Integral de Educación Digital, del Ministerio de Educación de la Nación, la declaración de importancia estratégica para la enseñanza y aprendizaje de la programación en el sistema educativo nacional durante la escolaridad obligatoria por parte del Consejo Federal de Educación (Res. 263/15 CFE) [8].

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.

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 generar entornos de programación que permitan desarrollar habilidades para la resolución de problemas, para que con su aplicación, los alumnos de diferentes niveles adquieran habilidades para desarrollar o potenciar el pensamiento computacional y superen el déficit en la capacidad de abstracción, recurriendo a estrategias que permiten una evolución del pensamiento abstracto partiendo de experiencias concretas.

Por otra parte, en la medida que el proyecto pueda alcanzar alumnos del nivel medio, estas acciones permitirán reducir los índices de deserción de los primeros años de las carreras de informática que se dictan en la UNTDF, contribuirán a aumentar la matrícula en las carreras vinculadas a la disciplina e incrementarán las capacidades de los alumnos para resolver problemas

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 1º 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.
c) Está en desarrollo, 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.
d) Se presentaron los avances del proyecto dentro de la Semana Nacional de la Ciencia y Tecnología, edición 2018, organizado por el Ministerio de Educación, Ciencia y Tecnología, Presidencia de la Nación, que permitió a alumnos de diferentes niveles educativos interactuar con los robots
e) Se están desarrollando talleres para alumnos de 1º y 2º 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.

f) Extender, al menos para alguno de los robots que se están utilizando, el conjunto de primitivas que el intérprete del robot reconoce.

g) Organizar talleres para alumnos del nivel medio, buscando despertar en los jóvenes la motivación necesaria para introducirse tempranamente en el mundo de la programación.

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, y dos (2) son alumnos avanzados de la carrera Licenciatura en Sistemas.

Los Licenciados Pendenti 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 defendido su tesis de grado, bajo la dirección de los docentes Depetris y Feierherd sobre temas afines al proyecto. (Aprobada con nota 10)

REFERENCIAS

[1] Academias, Conocimiento y Sociedad / Coordinadores Manuel A. Solanet ; Manuel L. Martí.- 1a Ed. compendiada. - Ciudad Autónoma de Buenos Aires: Academia Nacional de Ciencias Morales y Políticas, 2018. Pág. 59 ¨Cambio Tecnológico y la sociedad del futuro- Mario Solari - ISBN 978-987-99575-2-3

[2] Berrocoso J., Fernández Sánchez M. R., Garrido Arroyo M.; "El pensamiento computacional y las nuevas tecologías 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 10 de marzo de 2019.

[3] 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
Último acceso: 05/03/2019

[4] The New Zealand Curriculum on-line (s.f).
Recuperado de: http://nzcurriculum.tki.org.nz/The-New-Zealand-Curriculum/Technology
Último acceso: 11/03/2019.

[5]The Learn to Code
Recuperado de: https://www.codecademy.com/es
Último acceso 11/03/2019

[6] 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/diseniosccurriculares/primaria/2018/dis-curricular-PBA-completo.pdf
Último acceso: 09/03/2019.

[7] Queiruga C., Banchoff C. y otros (2018). "Escuelas TIC: las tecnologías digitales en las aulas". En proceedings del XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018)

Disponible:
http://sedici.unlp.edu.ar/handle/10915/67063.
Último acceso: 09/03/2019.
[8] Wing, Jannette M. (2009). Computational Thinking. Journal of Computing Sciences in Colleges vol.24, Issue 6 June 2009
Pages 6-7.
[9] Wing, Jannette M. (2008). Computational thinking and thinking about computing.
Philosophical Transactions of The Royal Society A, vol. 366, pág. 3717–3725.

Diseño y desarrollo de metodologías y herramientas para escenarios educativos híbridos

Sanz Cecilia[1,5], Madoz Cristina[1], Gorga Gladys[1], González Alejandro[1], Zangara Alejandra[1], Depetris Beatriz[2], Iglesias Luciano[1], Ibáñez Eduardo[1], Artola Verónica[1,4], Violini Lucía[1,3], Salazar Mesía Natalí[1,3], Archuby Federico[1,3], Sánchez Mariano[1], Nordio Mauricio[1], Fachal Adriana[1], Baldassarri Sandra[6], Manresa-Yee Cristina[7]

[1]Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC.
Facultad de Informática – Universidad Nacional de La Plata
[2]Universidad Nacional de Tierra del Fuego
[3]Becaria TIPO A UNLP
[4]Becaria Doctoral CONICET
[5]Investigador Asociado de la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)

[6]GIGA AffectiveLab, Universidad of Zaragoza, España

[7] Departamento de Ciencias Matemáticas e Informática. Universidad de las Islas Baleares, España

{csanz, cmadoz, ggorga, agonzalez, li, eibanez, vartola, lviolini, nsalazar, farchuby, msanchez, mnordio}@lidi.info.unlp.edu.ar,
alejandra.zangara@gmail.com, depetrisb@gmail.com, afachal@hotmail.com, sandra@unizar.es, cristina.manresa@uib.es

RESUMEN

En este trabajo se comparten las líneas de investigación y desarrollo en el marco del subproyecto "Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos", correspondiente al Instituto de Investigación en Informática LIDI. En particular, se presentan los resultados y avances realizados durante 2018 e inicios de 2019.
Se abordan 4 ejes temáticos principales, que a grandes rasgos, se vinculan con entornos virtuales, tecnologías digitales y metodologías para la colaboración, tendencias en paradigmas de interacción persona-ordenador para actividades educativas, y diseño y desarrollo de materiales y actividades educativas digitales. Como parte del proyecto, se participa en la formación de recursos humanos en el área a través de tesis de postgrado, becarios de investigación y proyectos con alumnos.

Palabras clave: Trabajo y aprendizaje colaborativo, Entornos digitales, Materiales educativos digitales, Objetos de aprendizaje.

CONTEXTO

Este subproyecto llamado "Metodologías y herramientas para la apropiación de tecnologías digitales en escenarios educativos híbridos" forma parte de un proyecto más general titulado: "Metodologías, técnicas y herramientas de Ingeniería de Software en escenarios híbridos. Mejora de proceso" (período 2018-2021), 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

El crecimiento acelerado de diferentes tecnologías tales como los dispositivos móviles, las redes que posibilitan la comunicación en todo momento y en todo lugar, y la disponibilidad de sensores, entre otros, ha dado lugar a escenarios donde se entrelazan el mundo real y el mundo digital. Es posible ver hoy situaciones donde se establecen, con inmediatez, hiperenlaces entre lo físico y lo digital (Aparici & García Marín, 2017; Liu, Dede, Huang & Richards, 2017). Cuando se escanea un código QR para acceder a más información de un producto, se crea un puente inmediato entre ambos mundos. Al mismo tiempo, los objetos del mundo real pueden mostrar efectos de la interacción de una persona en el mundo digital. El diseño y desarrollo de sistemas computacionales debe contemplar estos nuevos escenarios híbridos y distribuidos por lo que se requiere profundizar en la investigación en estas temáticas. Al mismo tiempo, estos escenarios modifican las prácticas cotidianas de las personas, sus hábitos y formas de interactuar. En el contexto educativo también se hacen presentes estas

hibridaciones, donde se combinan el entorno real y el virtual. Según Osorio (2010, pp.4): "*la concepción de lo híbrido desde la expansión, implica que el ambiente de aprendizaje, compuesto de actividades educativas, se desarrolla en situaciones presenciales, virtuales y autónomas, que se articulan estratégicamente buscando el enriquecimiento del proceso educativo […] Esto se concreta en el diseño de las actividades educativas, su desarrollo, evaluación y valoración.*"

El diseño de sistemas que integran paradigmas como la Realidad Aumentada, Interacción tangible y la Computación basada en dispositivos móviles y sensores, pueden favorecer la creación de escenarios educativos híbridos, que posibilitan recorridos que combinan el aula tradicional con el aprendizaje autónomo y extendido en ambientes digitales. La investigación en estas temáticas posibilita el real aprovechamiento de estas tecnologías. En este sentido, el subproyecto que aquí se presenta se focaliza en profundizar el estudio y desarrollo en esta área, por lo que en la próxima sección se presentan detalladamente las líneas de investigación y desarrollo en las que se trabaja. Cabe señalar que el subproyecto se vincula con las temáticas de la Maestría y Especialización en Tecnología Informática Aplicada en Educación de la misma Facultad, por lo que la sinergia entre ambos espacios favorece el desarrollo de tesis y trabajos finales en estos temas.

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.
- Paradigmas de interacción persona – ordenador: Realidad Aumentada (RA), Interacción Tangible sobre mesas interactivas, Computación ubicua, Realidad Virtual, desarrollo de Juegos Educativos con diferentes paradigmas de interacción.
- 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.

Cabe señalar que los distintos miembros del proyecto se organizan alrededor de las diferentes líneas de investigación del proyecto. Cada tema tiene un grupo de investigadores, becarios y tesistas que llevan adelante acciones de investigación y desarrollo. Al mismo tiempo, se realizan acciones de cooperación y vinculación con otras universidades del país y del exterior.

3. RESULTADOS OBTENIDOS

Se presentan aquí los principales resultados que se han alcanzado en el subproyecto, y corresponden al período 2018 - inicios de 2019.

Entornos digitales: campus virtuales, entornos virtuales de enseñanza y aprendizaje (EVEAs), entornos virtuales 3D, entornos simulados y laboratorios, redes sociales.

En esta línea se investigan tendencias en las arquitecturas, el diseño y la funcionalidad de entornos virtuales y su integración en escenarios educativos. También se investigan técnicas para aprovechamiento de la información y las dinámicas de las redes sociales para la creación de actividades educativas innovadoras. En cuanto a los entornos 3D se vienen analizando sus características y a partir de ellas las posibilidades y limitaciones para su uso en determinados contextos educativos. Desde el punto de vista teórico se vienen analizado trabajos que revisan la percepción de los entornos personales de aprendizaje por parte de los alumnos, y de los EVEA y campus virtuales, que son vistos como sistemas más institucionales (Dabbagh & Fake, 2017). Se analiza el rol de los usuarios como visitantes (solo consumen la información del entorno, lo usan como herramienta ocasional o institucional) o residentes (se sienten parte del entorno, participan activamente, producen contenidos) acorde a la clasificación de (Kuhn, 2017; White & Le Cornu, 2011). Se ha llevado adelante y finalizado la dirección de una tesis de maestría cercana al estudio de actitudes de docentes y alumnos frente a este tipo de tecnologías a partir de la cual se han realizado algunas publicaciones (Assinatto, Sanz, Gorga & Martin, 2018). También, en esta línea, en el marco del proyecto, se ha avanzado en el diseño y evolución del

entorno de enseñanza y aprendizaje IDEAS[1]. Se ha creado una comunidad del proyecto que acompaña su desarrollo[2]. Para ello, se trabajó con el repositorio *GitLab* y el gestor de contenedores *Docker*. Se agregaron funcionalidades de interés tales como la visualización de datos estadísticos que pueden orientar la toma de decisiones de los docentes y tutores del curso para el mejor acompañamiento de los estudiantes. También se avanzó con el desarrollo e integración de una herramienta de reuniones virtuales coordinadas que surge a partir de una tesina de grado de la Facultad (Digiani, Sanz & Gorga, 2018).
Se continuó con la integración de la herramienta de autoevaluación de IDEAS a un juego para dispositivos móviles (Desafiate) que permite presentar las preguntas de una evaluación con una serie de desafíos en el marco de una aventura pirata (Archuby, Sanz & Pesado, 2018). Actualmente se está trabajando en visualizar en IDEAS los desafíos que va ganando cada estudiante a través del juego.
Se realizó la segunda experiencia de actividad colaborativa[3], que integró el uso de *Twitter* y el EVEA IDEAS. Los alumnos eligieron sus propias herramientas de comunicación y producción para llevar adelante la tarea, por lo que generaron su propio ecosistema digital.
Respecto de la investigación sobre mundos virtuales 3D y entornos inmersivos, una integrante del proyecto participó en la dirección de dos tesis de maestría en estos temas, finalizadas en 2018. Se continuó con una tesis doctoral que indaga sobre las posibilidades de este tipo de entornos para personas con hipoacusia (Fachal, Sanz & Abásolo, 2018).

Diseño, producción y evaluación de Materiales educativos digitales. Objetos de aprendizaje y sus Repositorios.

Se continuó trabajando con la metodología CROA (Sanz, Barranquero & Astudillo, 2017; Sanz, Barranquero & Moralejo, 2016), que es utilizada en el marco de la Facultad de Informática y en particular, como parte de la práctica de un curso del doctorado. En 2018, se trabajó con dos objetos de aprendizaje desarrollados en el marco del proyecto. Uno sobre el concepto de autorregulación del aprendizaje, que se utilizó en el trabajo con escuelas de la región para llevar adelante acciones de articulación escuela media y universidad. Se trabajó también con los alumnos de primer año de la Facultad con el OA orientado a comprender el *merge* de listas ordenadas (Artola & Sanz, 2017) y con otro sobre compuertas lógicas (Estrebou, Salazar & Sanz, 2017). Además, en el marco de cursos de postgrado se inició el uso de otro OA desarrollado por integrantes del proyecto (http://163.10.22.82/OAS/ConceptoOA/).
También se diseñó un espacio para visibilizar los OA que se producen en el contexto de la Facultad y que están disponibles en una colección dentro del repositorio institucional de la universidad (SEDICI). El sitio está disponible en http://roa.info.unlp.edu.ar
Se avanzó con la investigación sobre sistemas recomendadores de objetos de aprendizaje y la evaluación de calidad de estos materiales educativos (Sanz & Barranquero, 2018).
Se está realizando una tesis doctoral con una becaria de la UNLP vinculada al diseño de un framework para el desarrollo de objetos de aprendizaje. Durante 2018 se realizó una revisión sistemática de literatura para profundizar en el estado del arte de la tesis en cuestión (Violini, Sanz & Pesado, 2018).
Se avanzó con la difusión de un trabajo de tesis de maestría en temas de materiales educativos hipermediales (Del Río, Sanz & Buccari, 2019).
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 y en su tesis de maestría de una becaria del equipo de investigación (Salazar, Sanz & Gorga, 2019). También se viene participando en estas temáticas a través de otros proyectos con universidades del país (Saldivia Obando, Gibelli & Sanz, 2018) y del exterior (Coma-Roselló, Aguelo-Arguis, Alvarez, Sanz & Baldassarri, 2018).

Trabajo colaborativo mediado por TICs.

En esta línea se estudian, diseñan y desarrollan herramientas y metodologías para mediar procesos colaborativos en contextos educativos. En 2018 se finalizó una tesis de doctorado (Zangara & Sanz, 2018) en relación a indicadores y herramientas de seguimiento del trabajo colaborativo. Por otra parte, se continúa con la investigación sobre la autorregulación del aprendizaje individual y grupal en procesos colaborativos a través de una tesis doctoral (Fierro, Sanz & Zangara, 2017) y un trabajo de especialización que está próximo a entregarse (Dieser, Sanz & Zangara, 2017).

Paradigmas de interacción persona – ordenador: Realidad Aumentada (RA), Interacción Tangible sobre mesas interactivas, Computación ubicua, Realidad Virtual. Desarrollo de Juegos

1 IDEAS es una evolución del EVEA WebUNLP y se encuentra en uso en la Facultad de Informática en diferentes cursos de grado y postgrado.

2 Proyecto IDEAS: http://proyectoideas.info.unlp.edu.ar

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.

Educativos con diferentes paradigmas de interacción

Sobre este eje se ha avanzado en varios aspectos durante 2018. Se estudiaron marcos teóricos sobre interacción tangible (IT), haciendo una revisión de 62 artículos de actualidad en el marco de una tesis doctoral sobre esta temática (Artola & Sanz, 2019). También en el marco de esta tesis se ha avanzado en el desarrollo de una herramienta de autor para la creación de actividades educativas basadas en interacción tangible. La tesis se desarrolla con una Beca CONICET.
Se ha desarrollado en el marco de proyectos de innovación con alumnos el juego El Conquistador (http://163.10.22.82/OAS/juegoconquistador/elconquistador2.mp4). Se trata de un juego de preguntas y respuestas como resultado de la investigación en temas de IT en escenarios educativos. Este juego se está evaluando con docentes y alumnos de escuelas de la región. Funciona sobre una mesa interactiva y se participa con dos avatares físicos que presentan diferentes tipos de *feedback* durante el juego.
También se avanzó en la difusión de FraccionAR, otro juego desarrollado basado en IT para el aprendizaje de fracciones. Éste fue presentado a docentes y ha sido evaluado con grupos de alumnos durante el 2018 (Sanz, Nordio & Artola, 2018). A lo largo del año se ha ido avanzando en la integración de alumnos para participar en temas de desarrollo de juegos IT.
Se está desarrollando una tesis de maestría vinculada a la investigación sobre Realidad Aumentada, y otra sobre Realidad Virtual (Chirinos & Sanz, 2018). Esta tesis se relaciona con un proyecto de extensión de la UNLP llamado "Huellas patrimoniales" que se propone acercar los patrimonios de la Argentina a niños (Comparato & Sanz, 2019). Se ha avanzado en el desarrollo de una aplicación de RV con microjuegos que permiten realizar un recorrido por las Cataratas del Iguazú. Se proyecta continuar con este trabajo en 2019.
Además, se está investigando sobre el tema juegos serios, que utilizan diferentes modelos de interacción persona-ordenador. En este sentido, se terminó un trabajo de especialización (Sandí Delgado & Sanz, 2018) y uno de maestría (Sandí Delgado, Lovos & Sanz, 2018).
Se avanzó en el uso de Astrocódigo con docentes y alumnos de escuelas de la región y de la Facultad de Informática, a partir del Proyecto Nexos y Proyecto Reforticca (Sanz, Artola, Miceli & Bione, 2018).
Además, se profundizó en la cooperación con otras universidad del país (Lovos, Gibelli y Sanz, 2017) y el exterior, en particular en temas de *m-learning* (Herrera, Manresa-Yee & Sanz, 2018).

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 2018 se participó en un proyecto de innovación docente de la Universidad de Zaragoza (PIIDUZ). Se realizaron investigaciones y publicaciones conjuntas. 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 y alumnos de grado que se vinculan en estas investigaciones.
En 2018, se han aprobado 7 trabajos de Maestría y 2 de Especialización en el área de Tecnología y Educación, dirigidos por miembros de este proyecto. También 1 tesis doctoral de un miembro del proyecto. Además se han finalizado 3 tesinas de grado dirigidas por miembros del proyecto en relación a los temas aquí presentados.

5. BIBLIOGRAFIA

- Aparici, R & García Marín, (coords.). ¡Sonríe te están puntuando! Narrativa digital interactiva en la era de Black Mirrow. Editorial Gedisa.
- Archuby, F; Sanz, C. & Pesado, P. (2018) Desafiate: un juego para la autoevaluación. Actas del XIII Congreso sobre Tecnología en Educación & Educación en Tecnología (TE&ET

2018), ISBN: 978-950-766-124-2, págs. 209-212, 2018. Disponible en: http://sedici.unlp.edu.ar/handle/10915/69073

- 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
- Artola, V. & Sanz, C. 2019). Informe técnico de avance de tesis doctoral sobre Marcos teóricos que sustentan la Interacción tangible en procesos educativos. III LIDI.
- Assinatto, G., Sanz, C., Gorga, G. & Martin, V. (2018). Actitudes y percepciones de docentes y estudiantes en relación a las TIC. Revisión de la literatura. Revista iberoamericana de tecnología en educación y educación en tecnología (ISSN 1850-9959), num. 22, pp. 7 -17. Disponible en: http://teyet-revista.info.unlp.edu.ar/TEyET/article/view/1166/935
- Chirinos, Y., Sanz, C. & Dapoto, S. (2018) La realidad virtual como mediadora de aprendizajes. Desarrollo de una aplicación móvil de realidad virtual orientada a niños. Propuesta de tesis de Maestría en Tecnología Informática Aplicada en Educación.
- Coma-Roselló,T.; Aguelo-Arguis, A.; Álvarez, P., Sanz, C. & Baldassarri, S. (2018). Analysis of Innovative Approaches in the Class Using Conceptual Maps and Considering the Learning Styles of Students," in *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 13, no. 4, pp. 120-129, Nov. 2018. doi: 10.1109/RITA.2018.2879388 URL: http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8532120&isnumber=8556148
- Comparato, G. & Sanz, C. (2019). Proyecto de Extensión aprobado por UNLP: Huellas Patrimoniales. Viajes inclusivos y nuevas experiencias sensoriales por medio de realidad virtual.
- 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.; Sanz, C. & Buccari, N. (2019) Incidence of a hypermedia educational material on the Teaching and Learning of Mathematics. En prensa. Journal New Approaches In Educational Research. Vol. 8. No. 1. January 2019. 2254-7339 DOI: 10.7821/naer.2019.1.334
- 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. (2018) *InfoMeeting*: una herramienta de comunicación sincrónica moderada. Aprobada en 2018. 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. (2018). Informe técnico sobre "Análisis de casos desarrollados en entornos virtuales 3D orientados 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.
- Herrera, S.; Sanz, C. & Manresa-Yee, C. (2018). Informe técnico de la estancia de investigación de la Dra. Susana Herrera en el marco de la cooperación con la Universidad de Islas Baleares y la Universidad Nacional de La Plata.
- 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
- Liu, D.; Dede, C.; Huang, R. & Richards, J. (editors) (2017). Virtual, Augmented, and Mixed Realities in Education. Springer.
- 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.

- Osorio G. Luz Adriana (2010). Ambientes híbridos de aprendizaje: elementos para su diseño e implementación. SISTEMAS: Ambientes Educativos Modernos Basados en Tecnología (117) (2010), pp. 70-79. http://www.acis.org.co/fileadmin/Revista_117/Uno.pdf
- Salazar, N.; Sanz, C. & Gorga, G. (2019). Análisis comparativo de librerías de realidad aumentada. Sus posibilidades para la creación de actividades educativas. Especialización en Tecnología Informática aplicada en Educación. Próxima a entregarse en 2019.
- Saldivia Obando, A., Gibelli, T. & Sanz, C. (2018). Propuesta pedagógica para la comprensión del espacio tridimensional utilizando Realidad Aumentada. Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018).
- Sandí Delgado, J.C; Sanz, C. (2018). Revisión y análisis sobre competencias tecnológicas esperadas en el profesorado en Iberoamérica. Revista electrónica de Tecnología Educativa. Edutec-e. (66), 93-121. ISBN 978-950-658-472-6
- Violini, L.; Sanz, C. & Pesado, P. (2018). Diseño, desarrollo, publicación y evaluación de Objetos de Aprendizaje. Un estudio de casos. Actas del XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018). ISBN: 978-950-658-472-6. Octubre de 2018.
- Sandí Delgado, J.C.; Sanz, C. & Lovos, E. (2018). Juegos serios para la indagación de competencias tecnológicas que puedan integrarse en la práctica pedagógica del profesorado. Una propuesta de aplicación en la Sede del Atlántico de la Universidad de Costa Rica (UCR). Tesis de Maestría en TIAE. Disponible en: http://sedici.unlp.edu.ar/handle/10915/71063
- Sanz, C.; Nordio, M.; & Artola, V. (2018). FraccionAR: Un juego para aprender sobre fracciones basado en interacción tangible. Actas del XIII Congreso sobre Tecnología en Educación & Educación en Tecnología (TE&ET 2018), ISBN: 978-950-766-124-2, págs. 84-92, 2018.
- Sanz, C. & Barranquero, F. (2018). 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
- Sanz, C.; Barranquero, F. & Astudillo, G. (2017) El concepto de OA bajo una mirada tecno-educativa. Debates y tendencias. Tutorial dictado en el marco de Laclo 2017. La Plata. 2017.
- Sanz, C.; Artola, V.; Miceli, P. & Bione, J. (2018) Astrocode in the wild, EDULEARN18 Proceedings, pp. 7542-7548. Pages: 7542-7548 Publication year: 2018 ISBN: 978-84-09-02709-5 ISSN: 2340-1117 doi: 10.21125/edulearn.2018.1759. Conference name: 10th International Conference on Education and New Learning Technologies. Palma, Spain
- Zangara, A. & Sanz, C. (2018) Interacción e Interactividad en el trabajo colaborativo mediado por tecnología informática. Tesis doctoral. Faculta de Informática. UNLP. Disponible en: http://sedici.unlp.edu.ar/handle/10915/67175
- White, D. & Le Cornu, A. (2011). Visitors and residents: a new typology for online engagement. *First Monday*, *16*(9), pp.1–10. doi:10.5210/fm.v16i9.3171

Diseño, implementación y evaluación de propuestas didácticas con realidad aumentada

Cecilia Sanz[2]; Tatiana Gibelli[1,3]; Edith Lovos[1]; Álvaro Saldivia; Sergio Condó; Paula Suárez[1]; Verónica Cuevas[3]

[1]Centro 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 incorporación de tecnologías emergentes en los procesos didácticos requiere de un cuidadoso y adecuado diseño, seguimiento y evaluación en función de los objetivos pedagógicos propuestos. Desde un proyecto de investigación acreditado por la Universidad Nacional de Río Negro (UNRN) se abordó este tema a través de una metodología de investigación-acción colaborativa. El objetivo del proyecto es investigar metodologías y estrategias innovadoras que favorezcan procesos educativos mediados por tecnologías digitales. Se puso énfasis en explorar tecnologías emergentes como la realidad aumentada, en particular con el uso de dispositivos móviles, y analizar sus posibilidades y barreras en procesos educativos. Se abordaron distintos aspectos tales como metodologías y herramientas para el diseño de actividades con RA, posibilidades de inclusión, y evaluación de estas propuestas. En este trabajo se presentan detalles de la investigación realizada en el marco de este proyecto que se encuentra en su etapa final.

Palabras clave: Tecnologías de la Información y la Comunicación, Realidad Aumentada, Propuestas Pedagógicas, Metodologías.

CONTEXTO

Esta investigación se desarrolla en el marco del proyecto "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, iniciado en 2016.

1. INTRODUCCIÓN

Las TIC atraviesan las actividades cotidianas y en especial han tamizado los contextos educativos. Las tecnologías digitales se caracterizan por permear el comportamiento de todos sus actores (estudiantes y docentes) en su vida cotidiana como sostiene (Silva, 2013). Desde la perspectiva docente, las TIC además de brindar la posibilidad de realizar ciertas actividades de manera más rápida, cómoda o fiable, permiten crear o complementar estrategias didácticas y, con ellas nuevos entornos de aprendizaje, y es aquí donde reside su verdadero potencial transformador (Cabero

Almenara, 2010). Sin embargo, la integración curricular de las TIC implica un proyecto educativo que implique un proceso de cambio e innovación educativa, de manera tal de lograr "la invisibilidad de la tecnología para una visibilidad del aprender" (Sánchez, 2002) .

En este sentido, el diseño de propuestas pedagógicas mediadas por TIC, en especial aquellas tecnologías consideradas emergentes como la RA, requieren un acompañamiento desde la investigación que permita dar sustento tanto desde marcos teóricos como desde las metodologías y herramientas que se ponen en juego. Asimismo es necesario un seguimiento adecuado de las experiencias en contexto que permitan evaluar tanto las tecnologías usadas como el logro de los objetivos pedagógicos.

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.

En relación a dichos objetivos las principales líneas de investigación abordadas en este proyecto fueron:

- Marcos teóricos sobre aprendizaje en contextos mediados por TIC.
- Realidad aumentada: exploración de usos, experiencias, recursos y herramientas.
- Diseño de propuestas didácticas que incorporen RA.
- Implementación y evaluación de propuestas pedagógica con RA.
- Metodologías de diseño de propuestas pedagógicas y materiales didácticos digitales con RA.
- Instrumentos para recolección de datos y evaluación de aprendizajes mediados por TIC.

3. RESULTADOS

Se avanzó en el proyecto acorde a las etapas planificadas y lo objetivos propuestos. Se siguió una metodología basada en un proceso de investigación-acción (Kemmis, Mc Taggart, 1998).

Se inició con una etapa de *diagnóstico*, que implicó delimitar los temas de investigación y realizar una revisión bibliográfica que incluyó temas como:

- Teorías de aprendizaje para contextos mediados por tecnologías digitales. En particular se profundizó en la Teoría Cognitiva del Aprendizaje Multimedia (Mayer, 2005).
- Tecnologías emergentes, en particular la realidad aumentada, y su uso en diferentes contextos educativos y con dispositivos móviles; focalizando en las áreas de matemáticas y programación (áreas de formación y espacios de práctica de docentes participantes en el proyecto).
- Aplicativos y herramientas para el diseño de propuestas didácticas con realidad aumentada. Se realizó un 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 (Lovos y Sanz, 2017).

- Exploración de herramientas de software disponibles para realizar modelos 3D que se insertarán en actividades de RA con énfasis en aquellas que permiten el trabajo con funciones matemáticas en tres dimensiones, profundizando en dos de ellas: Blender y SketchUp (Gibelli, Graziani, y Sanz, 2017).
- Revisión de materiales educativos digitales (MED) basados en RA para el aprendizaje de conceptos básicos de programación (estructuras de control y estructuras de datos dinámicas) en contexto universitario. Los MED explorados se pusieron a prueba en dos experiencias. Los resultados fueron aceptados para su presentación congreso (IPCTIIC 2017).
- Revisión de instrumentos y variables para evaluar materiales educativos digitales que incorporan realidad aumentada. En base a esta revisión (Lovos, Gibelli y Sanz, 2017) se diseñó un cuestionario que fue utilizado para recolección de datos en evaluación de implementaciones y en base a los resultados se espera validar el instrumento elaborado.

Posteriormente se llevó adelante la etapa *planificación*. En base a la revisión realizada se comenzó con el diseño de propuestas y materiales educativos con uso de realidad aumentada. En este sentido podemos mencionar el diseño de propuestas con uso de RA, a saber:

1. Actividad para el aprendizaje de conceptos de geometría tridimensional, destinada a estudiantes universitarios de primer año, basada en establecer asociaciones entre objetos cotidianos con conceptos de geometría tridimensional a través de la RA. Se hizo uso del aplicativo HP Reveal (www.hpreveal.com) que permitió generar "*Auras*" a través de los cuales se agrega información virtual a los objetos físicos considerados.
2. Diseño de un material educativo digital para la enseñanza de temas de cálculo, específicamente, la resolución de problemas de optimización. Se trata de una guía de estudio donde algunos problemas incluyen un código QR que enlaza con un sitio web de Geogebra donde se accede a una animación que ayuda a su comprensión. La actividad está destinada a estudiantes universitarios de segundo y tercer año.
3. Adecuación de las actividades del material educativo EPRA. Las actividades con RA se adecuaron para posibilitar su utilización en el marco de experiencias de enseñanza programación en la UNRN.

Luego en la etapa de *acción* se realizaron experiencias de diferentes propuestas didácticas con RA y el uso de dispositivos móviles. Durante estas experiencias se llevó adelante una recolección de datos que posibilitaron su análisis. Se siguieron metodologías estudiadas en la revisión bibliográfica con aportes propios del equipo del proyecto para la evaluación y diseño de las experiencias. En particular podemos mencionar como implementaciones de los diseños realizados las siguientes:

- Experiencias con la propuesta didáctica usando HP Reveal para comprensión de geometría tridimensional. Se realizó una experiencia piloto durante el segundo cuatrimestre de 2017 en la asignatura Matemática I correspondiente al primer año de la carrera de Licenciatura en Sistemas e Ingeniería Agronómica, y con ajustes en 2018. También se realizó una implementación con adaptaciones a la propuesta en la asignatura Álgebra de la carrera de Contador Público, durante el primer cuatrimestre de 2018 (Saldivia, Gibelli y Sanz, 2018).
- Experiencias de uso del material educativo digital con códigos QR para resolución de problemas de cálculo. Se llevó adelante una experiencia piloto en la asignatura Análisis Matemático correspondiente al tercer año de la carrera de Contador Público, durante el segundo cuatrimestre de 2017 y luego

una posterior implementación en la asignatura Matemática II de las carreras de Licenciatura en Sistemas e Ingeniería Agronómica, durante el primer cuatrimestre de 2018.

- Actividad de exploración y análisis de la experiencia de usuario (UX), de un material educativo digital basado en RA para el área de biología, con profesores e investigadores de las carreras de Ciencias del Ambiente, Ingeniería Agronómica y Lic. en Nutrición, que se dictan en la Sede Atlántica de la UNRN, durante segundo cuatrimestre de 2018

Finalmente, en una etapa de *reflexión,* que está en desarrollo, se analizan los resultados obtenidos en las diferentes experiencias. Esto implica una labor de síntesis de los principales resultados y elaboración de conclusiones. Se han realizado algunas comunicaciones de resultados parciales (Sanz et al, 2018) y se prevé la realización de nuevas publicaciones que den cuenta de las metodologías exploradas, las fortalezas y debilidades observadas así como recomendaciones y líneas de trabajo futuras. Se considera que esto constituirá un aporte en la temática y puede resultar de interés para investigadores y docentes que estén trabajando en propuestas educativas mediadas por tecnologías, en particular con realidad aumentada, aplicada en contextos educativos.

4. FORMACIÓN DE RECURSOS HUMANOS

El equipo de investigadores del proyecto 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 académica y de investigación de los integrantes, la dirección externa por parte de la Dra Cecilia Sanz, docente de la Facultad de Informática de la Universidad Nacional de la Plata, 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. 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 algunos miembros del proyecto han realizado cursos de postgrado del Doctorado en Ciencias Informáticas (UNLP) y otros vinculados a la temática del proyecto de acuerdo a intereses personales y formación específica.

También, 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 (UNLP). En este sentido se destaca la finalización de una tesis de maestría sobre juegos serios y su integración para el desarrollo de competencias tecnológicas en el profesorado (Sandí Delgado, Sanz, y Lovos, 2018). Asimismo se realizó asesoramiento a alumnos de grado a través de dos becas de inicio a la investigación UNRN (2017) y una beca CIN (convocatoria 2017).

Además se realizaron actividades de extensión y transferencia que consideramos son un aporte a la formación en estos temas, entre que se destacan:

- Proyecto de Extensión: "Orientación Pedagógica y Tecnológica para mejorar los procesos de enseñanza y aprendizaje en un Centro de Formación Profesional". En el marco de este proyecto se desarrolló un taller de capacitación destinado a docentes de centros de formación de oficios durante primer cuatrimestre de 2018.
- Desarrollo de un Programa de Trabajo Social titulado "Uso de tecnologías digitales como mediadoras de procesos de comunicación y educación en una escuela de educación secundaria", a cargo de dos alumnos de la Carrera de Licenciatura en Sistemas (2018).
- Diseño y desarrollo del taller "Comunicación y Noticias con

Realidad Aumentada" en agosto de 2018, a cargo de integrantes de este proyecto y docentes de la Lic. en comunicación social de la UNRN. Allí se exploraron las posibilidades de la RA en el campo de la comunicación y se realizaron prácticas con las herramientas Blippar y HPReveal.

- Dictado de taller "Exploración de herramientas y materiales educativos basados en Realidad Aumentada" en congreso Congreso Internacional: Humanidades Digitales, noviembre de 2018.

5. REFERENCIAS

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).

Kemmis, S. y Taggart, M. (1998). Robin: Cómo planificar la investigación-acción.

Lovos, E., & Sanz, C. (2017). Herramientas para Crear Actividades Educativas de Realidad Aumentada por Geolocalización. Análisis Comparativo. In Educación científica e inclusión sociodigital: actas del IX Congreso Iberoamericano de Educación Científica (CIEDUC 2017) (pp. 1470-1479).

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).

Mayer, R. E. (2005). The Cambridge Handbook of Multimedia Learning. New York: Cambridge University Press.

Saldivia, A., Gibelli, T. y Sanz, C.(2018). Propuesta pedagógica para la comprensión del espacio tridimensional utilizando Realidad Aumentada. In XXIV Congreso Argentino de Ciencias de la Computación (CACIC, Tandil, 2018).

Sánchez, J. (2002). Integración curricular de las TICs: conceptos e ideas. Actas VI Congreso Iberoamericano de Informática Educativa, RIBIE (pp. 20-22).

Sandí Delgado, J. C. (2018). Juegos serios para la indagación de competencias tecnológicas que puedan integrarse en la práctica pedagógica del profesorado: Una propuesta de aplicación en la Sede del Atlántico de la Universidad de Costa Rica (UCR). Tesis de Maestría en Tecnología Informática Aplicada en Educación. Universidad Nacional de la Plata.

Sanz, C. V., Gibelli, T. I., Lovos, E., Suárez, P., Saldivia, Á., Condó, S., & Cuevas, V. (2018). Realidad aumentada y otras tecnologías emergentes en procesos de enseñanza y aprendizaje: aproximaciones metodológicas al diseño y evaluación de propuestas didácticas. In XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).

Silva, J. (2013). TIC en educación superior. Una reflexión teórica-práctica. En Revista de la Asociación Colombiana de Ciencias(Col.), 25: 92-97.

Educación a Distancia en el ámbito de la UNSAdA

Claudia Russo[1], Nicolás Alonso[1], Mónica Sarobe[1], Sebastían Mir[1], María Celeste Nobal[2], Eliana Serrano[1], Marina Lilian Rodríguez[1], Mario Gabriel Ábalos[3], Matías Contreras[3], Gabriel Sosa[3], Ezequiel Stutz Diaz[3], Patricio Belesia[3]

Universidad Nacional de San Antonio de Areco (UNSAdA)

Avenida Güiraldes N° 689 (CP 2760, San Antonio de Areco, Bs.As.) - Teléfono: (02326) 453079

crusso@docentes.unsada.edu.ar, nfalonso@docentes.unsada.edu.ar, monicasarobe@unnoba.edu.ar, sebastian.mir@unsada.edu.ar, celeste.nobal@unsada.edu.ar, elianaserrano@unnoba.edu.ar, mrodriguez@docentes.unsada.edu.ar, {mgaabalos, mjcontrera, gasosa, pbelesia, estutzdiaz}@alumnos.unsada.edu.ar

Resumen

Se presenta un proyecto de *I+D+i* (Investigación, Desarrollo e innovación) centrado en las tecnologías educativas y su relación con el contexto de la Educación a Distancia (EaD). En el marco del presente proyecto y considerando que las tecnologías actuales plantean nuevos paradigmas que implican un impacto directo en los modelos educativos, se investigará sobre temas relacionados con las tecnologías informáticas aplicadas a la educación, haciendo énfasis en las diversas variantes y evoluciones de la modalidad de EaD.

Se indagará también sobre distintas metodologías y técnicas que permitan dotar de interoperabilidad a los sistemas de información presentes en la Universidad Nacional de San Antonio de Areco (UNSAdA), que tengan relación con la modalidad de EaD, realizando experiencias concretas.

Para ello se propone construir, mediante herramientas de software de código abierto, un Entorno Virtual de Enseñanza y Aprendizaje (EVEA) que integre dichos sistemas de información para lograr una administración y gestión académica de forma centralizada que permita crear, administrar y llevar a cabo aulas virtuales, herramientas de aprendizaje, comunicación y colaboración.

Acompañará a la presente investigación la generación de una guía que posibilite la producción, administración y almacenamiento de contenidos digitales bajo la forma de Recursos Educativos Digitales (RED). La guía se propondrá desde una mirada interdisciplinaria que permita determinar y sistematizar aspectos determinantes en el desarrollo de materiales didácticos digitales.

Los resultados del presente proyecto sentarán precedentes en temas de la modalidad de EaD que podrán incluirse en un eventual armado del Proyecto Institucional de la UNSAdA frente al Sistema Institucional de Educación a Distancia (SIED) dispuesto el Ministerio de Educación y Deportes de la Nación, que reglamenta y fija criterios sobre la modalidad estudiada.

[1] Docente Investigador de la Escuela de Desarrollo Productivo y Tecnológico | UNSAdA
[2] Personal Administrativo No Docente | UNSAdA
[3] Becario Alumno | UNSAdA

Palabras clave: Entornos Virtuales, Educación a Distancia, Sistema Institucional, Recursos Digitales, Enseñanza-Aprendizaje.

Contexto

La Universidad Nacional de San Antonio de Areco (UNSAdA) fue creada el 25 de noviembre del año 2015 por *Ley del Congreso de la Nación N° 27.213/2015*.

Se rige por las leyes nacionales, su ley de creación, su Estatuto y las reglamentaciones que en su consecuencia se dicten. Tiene su sede en *Avenida Güiraldes N°689* de la ciudad de San Antonio de Areco, Provincia de Buenos Aires, donde se encuentra el asiento de sus autoridades centrales.

Tiene como principal misión constituirse en un centro de referencia para el desarrollo del conocimiento científico-tecnológico estratégico y la generación de innovación en la región.

El proyecto que aquí se describe fue aprobado por *Resolución N°110/2018* del Consejo Superior en el marco de la Convocatoria *NACT 2017 | Formación de Núcleos de Organización y Ejecución de Actividades Científicas y Tecnológicas – NACT*.

El proyecto está integrado por un equipo interdisciplinario de investigadores, docentes y estudiantes pertenecientes a la Escuela de Desarrollo Productivo y Tecnológico, y se trabajará en conjunto con la Escuela de Desarrollo Social y Humano de la UNSAdA. Ambas Escuelas constituyen el ámbito curricular de la organización universitaria.

Introducción

Las tecnologías se presentan cada vez más como una necesidad en el contexto de la sociedad donde los rápidos cambios, el aumento de los conocimientos y las demandas de una educación de alto nivel y actualizada se convierten en una exigencia permanente.

La educación es un fenómeno social y como tal es necesario tener presente el objetivo primordial, que es contribuir al desarrollo integral o promoción de la persona, teniendo en cuenta los aspectos que van a facilitarle la inclusión en una sociedad determinada [1].

Es evidente que no es posible lograr este objetivo si no se considera a la sociedad en su conjunto, sus constantes transformaciones y a sus individuos y sus necesidades.

Las universidades de todo el mundo están impulsando la alfabetización digital, por considerarla un objetivo esencial en la formación de los futuros profesionales. En este sentido, es necesario tener presente las transformaciones que como sociedad hemos enfrentado en las últimas décadas, algunas de las cuales persisten hoy en día. Una de estas transformaciones, que continúa motivando cambios en nuestros hábitos y costumbres, ha tenido lugar gracias a la incorporación de las Tecnologías de la Información y la Comunicación (TIC), y al fenómeno de Internet, que ha transformado los procesos de comunicación y el acceso a la información y el conocimiento.

La sociedad actual está inmersa en un mundo de productos tecnológicos, este hecho está directamente relacionado con la revolución científico-técnica iniciada en el siglo XX y que se profundiza en el siglo XXI [2].

Los procesos de enseñanza y aprendizaje se han hecho más complejos, en parte gracias al auge de nuevas tecnologías, especialmente de internet. La potencialidad de estas tecnologías ha despertado el interés de docentes e investigadores por conocer y profundizar las

condiciones y características de nuevas formas de enseñar y aprender. Se observa en el contexto actual de educación superior que los docentes están optando por la modalidad de aprendizaje combinado como estrategia para incorporar la virtualidad en la propuesta de enseñanza [3]. Esta nueva forma de enseñar y aprender permite además la desaparición de las condiciones espacio-temporales y las posibilidades de conectarse en espacios dispersos geográficamente.

Todo acto educativo implica acciones comunicativas entre docente y estudiantes, quienes comparten información y la procesan para generar conocimiento. En el aula de clase, actividades como la exposición y discusión oral, la lectura de textos impresos, la ejercitación y la práctica en laboratorio se apoyan con materiales educativos como tablero, libros, documentos y manuales impresos. Estos materiales sirven como mediadores en el proceso enseñanza – aprendizaje, para comunicar los contenidos y facilitar su comprensión y apropiación.

Con las TIC es posible producir medios integrando texto, imagen, audio, animación, video, voz grabada y elementos de software, almacenarlos en computadores o llevarlos a Internet para ser leídos desde un computador o un dispositivo móvil. A estos medios se les conoce como medios digitales por el hecho de estar representados en un lenguaje binario, propio de las computadoras.

La incorporación de la EaD ha sido un proceso muy heterogéneo en las instituciones nacionales. Las estrategias implementadas nos hablan de instituciones que incorporan la gestión de la EaD en su organigrama y otras que solo avanzan en la integración de aulas virtuales como apoyo a las clases presenciales sin estar reglamentadas o encuadradas formalmente.

De allí que los desafíos, que atraviesan cada una de las universidades, no son los mismos. No obstante, ello, se comparte el desafío de gestionar buenas prácticas de enseñanza y de aprendizaje mediadas por las tecnologías que orienten a la formación de calidad de profesionales y de ciudadanos comprometidos socialmente.

Coherente con ello, desde la UNESCO en su informe mundial de la educación [4], se enfatiza la creación de entornos pedagógicos-tecnológicos, capaces de salvar las distancias y establecer sistemas de educación de alta calidad, favoreciendo así el progreso social, económico y la democratización; aprovechando plenamente las tecnologías de la información y la comunicación con fines educativos.

Es por este motivo que el uso de las TIC en procesos de formación exige considerar una serie de cuestiones y desafíos que se presentan.

A nivel administrativo, la creación de espacios de enseñanza y aprendizaje requiere una gestión de los recursos que se ven involucrados, no sólo recursos humanos sino también materiales y tecnológicos y, además, cuestiones de índole académico-administrativas vinculadas a la autorización para utilizar y/o administrar un espacio formativo en el ámbito universitario.

También se deben atender, con un nivel apropiado de calidad, las consultas que se reciban en lo referente a temas administrativos, pedagógicos y técnicos.

Asimismo, las tecnologías deben ser utilizadas criteriosamente y con un objetivo pedagógico claro y pertinente, esto sumado a que las mismas se encuentran en un proceso de continua evolución, pone de manifiesto la

necesidad de proveer mecanismos de capacitación y de acompañamiento al servicio de quienes serán sus usuarios finales.

Por último, hay que considerar el componente técnico que se ve involucrado al utilizar tecnologías, lo que supone atender a posibles inconvenientes técnicos y desafíos que podrían presentarse relacionados con la usabilidad de herramientas que, en principio, podrían resultar desconocidas.

Por lo anteriormente expuesto, el presente proyecto plantea diferentes líneas de investigación que tienen como meta el mejoramiento de los procesos formativos a través de la incorporación del uso de las TIC, abordando aspectos tales como el desarrollo y utilización de Entornos Virtuales de Enseñanza y Aprendizaje y su interconexión con los sistemas de información presentes en la UNSAdA en lo que a temas de EaD refieren.

Acompañan a este meta, el sentar los precedentes del conjunto de acciones, normas, procesos, equipamiento, recursos humanos y didácticos que permitirán el desarrollo de propuestas a distancia, como así también, la generación de una guía que posibilite la producción, administración y almacenamiento de contenidos digitales bajo la forma de Recursos Educativos Digitales (RED). La cual contendrá la sistematización de aspectos que facilitan la producción de materiales didácticos digitales, aspectos que favorezcan a la gestión y administración de dichos materiales.

Líneas de investigación y desarrollo

El actual proyecto se centrará en tres líneas de investigación y desarrollo principales, planteándose para cada una de ellas una serie de objetivos específicos, metas y resultados esperados:

Entornos Virtuales de Enseñanza y Aprendizaje:

- Investigar los sistemas de información presentes en el ámbito institucional y académico.
- Analizar y comparar los principales sistemas LMS (*Learning Management System*) de código abierto tomado como experiencia las utilizadas ya por otras Universidades.
- Analizar y comparar protocolos de acceso a servicios de red e inicio de sesión único, tales como: SSO (*Single Sign-On*) y LDAP (*Lightweight Directory Access Protocol*).
- Customizar, adaptar e implantar un sistema LMS, permitiendo la gestión de cursos, materiales y herramientas de aprendizaje, comunicación y colaboración.
- Implementar un sistema que se integre con la plataforma LMS para llevar a cabo clases a distancia de manera síncrona con la capacidad para generar y reproducir audio, video, chat y videoconferencia, entre otras.
- Crear instrumentos de medición que permitan relevar y evaluar la usabilidad del sistema LMS implantado, con la finalidad de poder recaudar información para futuros desarrollos.

Recursos Educativos Digitales

- Sistematizar un corpus de buenas prácticas en la producción, gestión y almacenamiento de materiales didácticos digitales.
- Indagar en las diferentes herramientas de la actualidad que suponen un cambio sustantivo en la relación sujeto-objeto del conocimiento.

- Analizar y comparar diferentes estándares relacionados con la creación de Recursos Digitales Educativos (RED) y otros recursos digitales similares usados para el apoyo al aprendizaje.

- Seleccionar un portfolio de aplicaciones de EaD que ofrezcan funcionalidades específicas y cubran metodologías pedagógicas concretas.

- Implementar materiales educativos digitales apoyándose en el uso de objetos de aprendizaje, multimedia e hipermedia, y aplicarlos en los diferentes niveles educativos de la UNSAdA, incorporando diferentes metodologías de diseño estudiadas en el marco del proyecto.

Sistema Institucional de Educación a Distancia

- Relevar las normativas, resoluciones, leyes y estándares, tanto a nivel nacional como internacional, en lo referente a los sistemas de Educación a Distancia (EaD) en educación superior.

- Recopilar y analizar bibliografía, revistas y publicaciones para determinar el estado del arte de la implementación de sistemas de EaD en Universidades Nacionales.

- Dominar conceptos y adquirir capacitación en las normativas relacionadas al Sistema Institucional de Educación a Distancia (SIED) dispuesto por el Ministerio de Educación y Deportes de la Nación.

- Proponer roles, funciones, acciones y procesos junto con el eventual equipamiento y gestión de recursos humanos y didácticos, que posibiliten el desarrollo de propuestas a distancia en el ámbito de la UNSAdA

Resultados y Objetivos

Se espera que los aspectos investigados contribuyan a propiciar el fortalecimiento en la formación de recursos humanos, en su rol de investigadores o partícipes activos en equipos de investigación; fomentando la culminación de sus estudios superiores, promoviendo la redacción, exposición y defensa de trabajos finales de grado y postgrado.

En lo que refiere a la implantación de un Entorno Virtual de Enseñanza y Aprendizaje (EVEA), se espera que los aspectos investigados contribuyan al fortalecimiento de las experiencias de enseñanza y aprendizaje llevadas a cabo por nuestra universidad.

En relación a los Recursos Digitales Educativos (RED), se espera constituir información básica para los actores intervinientes en el proceso de producción y gestión de los materiales digitales permitiendo una sistematización que facilite la producción, reutilización y administración de los mismos.

Como contribución institucional se pretende que los resultados del presente proyecto sienten precedentes en temas de Educación a Distancia (EaD) que puedan incluirse en un eventual armado del Proyecto Institucional de la UNSAdA frente al Sistema Institucional de Educación a Distancia (SIED) dispuesto el Ministerio de Educación y Deportes de la Nación, que reglamenta y fija criterios sobre la modalidad de EaD.

Formación de Recursos Humanos

El equipo está conformado por 2 (dos) investigadores formados, 3 (tres) investigadores en formación que desarrollan actividades de docencia e investigación dentro de la Universidad, junto a 1 (un) personal administrativo y 6 (seis) estudiantes avanzados de las carreras de Licenciatura en Informática.

Además, se realizan convocatorias abiertas a estudiantes de las carreras de las Escuelas de Desarrollo Productivo y Tecnológico, junto a la de Desarrollo Social y Humano para que se inicien en tareas de investigación.

El desarrollo de esta línea de investigación espera contribuir a generar en el plazo de dos años, tres Trabajos Finales vinculados al área de Tecnología y Educación que sean fuente de la obtención del título de Licenciatura en Informática. Dichos Trabajos Finales serán dirigidos por miembros del presente proyecto.

Referencias

[1] Aran A., Freixes N., Miravalles A. La educación social. Una mirada didáctica. Primera Edición. Editorial Graó. España. 2008

[2] Patricia Halaban. "La comunicación virtual en educación a distancia". Un estudio en interacciones comunicacionales y procesos pedagógicos en Internet". Ediciones CICCUS. 2010

[3] Stella Briones. Delegados y Horizontes para el siglo XXI. Problemas y desafíos siempre vigentes de la enseñanza y aprendizaje a distancia en nuestros contextos institucionales. RUEDA 2011.

[4] Informe Mundial de la Educación. UNESCO 1998. http://www.unesco.org/education/educprog/wche/declaration_spa.htm

Bibliografía

- F. García Córdova y M. R. Trejo, La perspectiva de la investigación tecnológica en educación. México: Limusa, 2012.
- Cabero, J., y García, F. (coords.), Realidad aumentada. Tecnología para la formación. Madrid: Síntesis. 2016.
- Casamayor, G.; Alós, M.; y otros. La Formación On-line, Una mirada integral sobre el e-learning, b-learning. Primera edición. Editorial Graó. Barcelona, España. 2008.
- Krumm J., Creating a New Book of Tutorials: Ubiquitous Computing Fundamentals, Microsoft Research Redmond, Ubicomp 2009.
- Lion, C. Imaginar con tecnologías, Relaciones entre tecnologías y conocimiento. Primera edición. Editorial Stella. Buenos Aires, Argentina. 2006.
- Barberà, E. La incógnita de la Educación a Distancia. Primera edición. Editorial Horsori. Barcelona, España. 2001.
- Litwin, E.; Maggio, M.; Lipsman, M. Tecnologías en las aulas, Las nuevas tecnologías en las prácticas de la enseñanza. Casos para el análisis. Primera edición. Editorial Amorrortu. Buenos Aires, Argentina. 2005.
- Litwin, E. Tecnologías educativas en tiempos de Internet. Primera edición. Editorial Amorrortu. Buenos Aires, Argentina. 2009.
- Manzanedo, J. El e-Learning en España, Modelos actuales y tendencias de actuación. Primera edición. Colección EOI Tecnología e Innovación. España. 2003.
- Almenara, J.; Graván, P. E-actividades, Un referente básico para la formación en Internet. Primera edición. Editorial MAD. Sevilla, España. 2008.

El Video para generar oportunidades de Enseñanza- Aprendizaje en Entornos Virtuales Flexibles

Alejandra Orellana; Cristina Vera

Departamento de Informática / FCEFN/ UNSJ
Av. Ignacio de la Roza 590 (O) - San Juan- Argentina
Telefonos:4260353 -4260355
ale_ore@yahoo.com.ar; civera2@yahoo.com.ar

Resumen: Seguir en clase la explicación del docente desarrollando un algoritmo, no garantiza la adecuada generación de competencias en el alumno, esta es una de las dificultades que surgen a la hora de enseñar programación. No obstante, hoy en día encontramos un gran número de herramientas y recursos que permiten trabajar con Entornos Virtuales Flexibles (EVF), éstos favorecen la consideración de la dimensión individual, las características únicas de cada alumno, con sus propias limitaciones, fortalezas, intereses, capacidades y estilos de aprendizaje además de presentar una alternativa válida para superar los limitantes encontrados en el uso de las plataformas institucionales. Este artículo expone brevemente una de las líneas de investigación relacionadas al diseño de micro-contenidos en formato de videos para ser utilizados en dichos entornos y así apoyar los desafíos que presenta el proceso de enseñanza-aprendizaje de la programación tanto para alumnos como para docente.

Palabras claves: *Video, entorno virtual flexible.*

Contexto

Proyecto de Investigación "*Los entornos virtuales flexibles para el desarrollo de competencias en el área de algoritmos y lenguajes de programación*" Código 21/E1113, vigente en el periodo 2018-2019, de la Universidad Nacional de San Juan. Argentina CICITCA. Resolución N°21/18-CS, el mismo forma parte del trabajo conjunto del Programa permanente de Investigación del Gabinete de Tecnología e Innovación Educativa (GATIE) y el Departamento de Informática. Este se encuentra estrechamente relacionado con el desarrollo de competencias de los alumnos de las carreras de Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información de la Universidad Nacional de San Juan.

1. Introducción

La construcción de algoritmos computacionales precisa de un nivel complejo de abstracción ya que requiere pensamientos de orden superior, tal cual se plantea en la Taxonomía de Bloom. Habitualmente los alumnos se inician con tareas básicas que incluyen recordar datos, cifras, conceptos y después se sigue un proceso más o menos estructurado para llegar al siguiente nivel: la comprensión de la información, después se avanza en el nivel de aplicación de nuevos procedimientos o estrategias, para luego avanzar al nivel de análisis para entender sus partes, más tarde avanzar al nivel de evaluar la información para luego llegar al nivel más complejo que consiste en crear un nuevo producto o un nuevo proceso desde un nuevo punto de vista sobre la base de la información original. Este último nivel es donde llegan los alumnos que obtuvieron mayor preparación académica (Amo, D, Santiago, R., 2017).

Habitualmente cuando un docente enseña programación, presenta la clase y los conceptos que conllevan, evacúa las dudas de

los alumnos referidas a definiciones y luego desarrolla los ejercicios en el pizarrón. Esta forma de enseñar a programar aspira a que los alumnos sean capaces de asociar las técnicas que usa el profesor cuando desarrolla el ejercicio en el pizarrón, para que luego, el alumno, pueda realizar esta asociación patrón-técnica en la resolución de nuevos ejercicios. Este tipo de aprendizaje, se puede llamar aprendizaje por imitación, ya que se basa en que el alumno sea capaz de generar las habilidades necesarias tratando de imitar lo que el docente impartió. Esto explica una buena parte de las dificultades de enseñar a programar y algunos de los siguientes problemas:

- Seguir en la clase la explicación del docente desarrollando un algoritmo, no garantiza la adecuada generación de capacidades en el estudiante, y el éxito depende de diversos factores que el docente no puede controlar. Esto hace que a ciertos alumnos les vaya bien, y a "otros" no tanto, sin que los docentes puedan hacer nada al respecto.

  •El proceso de construcción de un algoritmo conlleva varias maneras de encararlo y por lo tanto da la sensación que el "cómo hacer las cosas" no es enseñable, y que es algo que depende del ingenio del que está construyendo el algoritmo (Villalobos, J. A., & Calderón, N. A. , 2009)

Estas son algunas de las dificultades encontradas en la enseñanza de la programación, por ello en la actualidad, tanto docentes como alumnos utilizan videos para distintos propósitos, especialmente para enseñar y aprender y reforzar la explicación del docente.

Investigadores en Canadá encuestaron una muestra de 2000 participantes y estudiaron la actividad cerebral de otros 112 usando electroencefalogramas (EEG), esta información fue usada por Microsoft, que descubrió que desde aproximadamente el año 2000 cuando comenzó la revolución de los dispositivos móviles, las personas pierden la concentración después de ocho segundos, esto también se relaciona con algunos comportamientos o hábitos, por ejemplo cuando la atención de la persona no está ocupada por ninguna otra actividad, buscan algo en su dispositivos móviles (McSpadden, K., 2015).

Este equipo de investigación, considera que el vídeo es uno de los recursos que puede ser explotado en el ámbito educativo debido al impacto audiovisual, y el modo asíncrono de ser visto, como así también el ser replicado tantas veces como sea necesario, permitiendo presentar de forma visual y auditiva el paso a paso de la realización de una actividad, la explicación de un contenido o procedimiento, y por otro, posibilita que el alumno aprenda de manera autónoma, resuelva dudas y avance en el aprendizaje a su propio ritmo. Por ello, parte del equipo de este proyecto de I/D participó de la convocatoria de "COOL VIDEOS", diseñado y organizado por la Universidad Nacional del Noroeste de la Provincia de Buenos Aires. para la *"formación de docentes en la generación de materiales educativos digitales"*, formando parte de la capacitación para el diseño de videos educativos. Esta capacitación no solo permitió profesionalizar la creación de contenidos audiovisuales, sino que también generó el desafío de trabajar con micro-contenidos para ser visualizados en dispositivos móviles, donde este último forma parte de los temas de Tesis de Maestría incluida en este Proyecto.

El video constituye un factor crítico para lograr resultados de aprendizaje; y es una herramienta eficaz para la enseñanza y el aprendizaje en diversas disciplinas. Sin embargo, simplemente presentando la información en formato de video no conducirá automáticamente a un aprendizaje en profundidad (Karppinen P. ,2005). El diseño pedagógico y el desarrollo de videos con elementos críticos es crucial para que sea una herramienta efectiva en contextos educativos (Gedera, D., & Zalipour, A., 2018).

Entorno Virtual Flexible

Los 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, P. R., García-Peñalvo, F. J., & Conde-González, M. Á., 2015). El uso de videos en dispositivos móviles constituye un desafío a la hora de diseñar micro-contenidos para que puedan correr sobre los mismos.

Clasificación de videos

Es posible distinguir distintos tipos de video, a saber: Pack de bienvenida, Refuerzo, Tutorial, Lección aprendida y tutoría proactiva (García-Peñalvo, F. J., 2018).

Tomando como base las encuestas ya realizadas sobre las herramientas web 2.0 que los alumnos utilizan con mayor frecuencia en su vida diaria (Llarena, M. G., Díaz, M., & Vera, C., 2017), también resulta útil indagar sobre las preferencias que tienen los alumnos sobre las tipologías de video existentes.

Así mismos, es importante investigar la importancia de diseñar estos contenidos para dispositivos móviles con criterios de calidad del video, las dificultades que se presentan respecto a la restricción por tamaño de las pantallas de los mismos, tiempo de duración acorde a explicar un tema puntual. Y la posibilidad de subirlo a un entorno LMS (Learning Management System) como Moodle o Schoology, o distribuirlo por medios informales de trabajo para su réplica.

Guion del video

Uno de los elementos más importantes a considerar al elaborar un vídeo didáctico es el guión del vídeo porque es aquí donde entra el detalle de la información que será transmitida, por un lado, se tiene la información relacionada con el concepto y por el otro los elementos adicionales que se incorporarán como son las imágenes, sonido, transición del vídeo, etc., lo que dará lugar a un medio audiovisual. El proceso del guión conlleva el conocimiento del concepto a transmitir, la creatividad e imaginación de quien pretende elaborarlo (Ramos, L. M., & Flores, T. G., 2014).

2. Líneas de investigación y desarrollo

La línea de investigación de este proyecto es proponer estrategias didáctico pedagógicas que integren el entorno institucional -asociado a la Educación Formal - con el conjunto de aplicaciones y servicios de la web 2.0 asociado al aprendizaje informal de los alumnos, y atendiendo conceptos 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 esta 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 una herramienta valiosa 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 los aspectos emocionales-afectivos a tener en cuenta.

En el presente artículo, se ha expuesto una de las líneas de I/D referidas al diseño de micro-

contenidos en formato de videos para ser utilizados en dichos entornos para así apoyar los desafíos que presenta la enseñanza de la programación tanto para alumnos como para docentes.

3. Resultados Esperados

En este apartado se muestran algunos de los objetivos que se ha planteado este proyecto de I/D algunos ya alcanzados y otros que están en curso.

Objetivos ya alcanzados:

- Identificar las competencias específicas para el aprendizaje de la Programación.
- Definir estrategias en el uso de un Entorno Virtual Flexible para el desarrollo de competencias para el aprendizaje de la Programación.
- 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.

El equipo de I/D ha estudiado diversas herramientas y ha llevado a cabo distintas experiencias en la asignatura Programación Procedural, que corresponde al primer año de las carreras Licenciatura en Ciencias de la Computación y Licenciatura en Sistemas de Información de la Facultad de Ciencias Exactas de la Universidad Nacional de San Juan (Llarena, M. G., & Díaz, M., 2018), (Vera, C., Llarena, M. G., & Díaz, M., 2017).

De esta experimentación se ha obtenido conclusiones para mejorar el uso de estas herramientas en distintos momentos de la clase o fuera de ella.

Algunos de los objetivos a concretar sobre la línea de investigación específica para este artículo son:

-Diseñar el material didáctico en formato de micro-contenidos para dispositivos móviles.

-Experimentar su uso en alumnos que cursan materias del área Algoritmos y Lenguajes de Programación,

-Evaluar el trabajo de campo realizado y las estrategias utilizadas.

4. Formación de recursos humanos

Cabe destacar que existen tres tesis de Posgrado enmarcadas en este Proyecto, donde una de ellas consiste en establecer estrategias para diseñar recursos didácticos para dispositivos móviles, y evaluar su uso e impacto.

5. Bibliografía

Amo, D., & Santiago, R. (2017). Learning analytics: la narración del aprendizaje a través de los datos. UOC.

García-Peñalvo, F. J. (2018). Contexto de investigación. Grupo GRIAL. Material curso MOOC: Flip Teaching: una metodología activa. Universidad Politécnica de Madrid.

Gedera, D., & Zalipour, A. (2018). Use of interactive video for teaching and learning. In Australasian Society for Computers in Learning in Tertiary Education, ASCILITE) (pp. 362-367). ASCILITE 2018 Deakin University.

Humanante-Ramos, P. R., García-Peñalvo, F. J., & Conde-González, M. Á. (2015). Personal learning environments and online classrooms: An experience with university students. IEEE Revista Iberoamericana de tecnologías del aprendizaje, 10(1), 26-32.

Karppinen, P. (2005). Meaningful learning with digital and online videos: Theoretical perspectives. *AACE journal*, *13*(3), 233-250.

Llarena, M. G., & Díaz, M. (2018). Los entornos virtuales flexibles para el desarrollo de competencias en el área de algoritmos y lenguajes de programación. In *XX Workshop de Investigadores en*

Ciencias de la Computación (WICC 2018, Universidad Nacional del Nordeste).

Llarena, M. G., Díaz, M., & Vera, C., (2017). Entornos Personales de Aprendizaje para la mejora de la calidad de blend-e-learning: Experiencia en un curso de 1er Año. IIV Congreso Virtual Iberoamericano de Calidad en Educación Virtual a Distancia, (Eduq@2017).

McSpadden, K. (2015). You Now Have a Shorter Attention Span Than a Goldfish. *Time.*

Ramos, L. M., & Flores, T. G. (2014). El vídeo como recurso didáctico para reforzar el conocimiento. Memorias del Encuentro Internacional de Educación a Distancia, (3).

Vera, C., Llarena, M. G., & Díaz, M. (2017). Entornos personales como apoyo para el aprendizaje y desarrollo de competencias: experiencia con alumnos de informática. In *XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET, La Matanza 2017).*

Villalobos, J. A., & Calderón, N. A. (2009). Proyecto Cupi2: un enfoque multidimensional frente al problema de enseñar y aprender a programar. *Revista de Investigaciones UNAD*, *8*(2), 45-64.

Empleo de ChatBots educativos como recurso complementario en las prácticas docentes

José Federico Medrano, Mario Alberto Tejerina, César Alejandro Castillo

jfmedrano@fi.unju.edu.ar, mariotejerina@gmail.com, ce_al_castillo@yahoo.com.ar,

VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería

Universidad Nacional de Jujuy - Ítalo Palanca 10, +54 (388) 4221587

RESUMEN

Los entornos virtuales se ofrecen como claros complementos a la labor docente, con los avances tecnológicos, la inclusión de estos dentro del aula ha propiciado un cambio y una revolución digital por más de una década. Estos avances no sólo han cambiado el modo de enseñar de los docentes sino también el modo de aprender de los alumnos. El aula ya no se encuentra definida por el espacio físico, la virtualidad expande los límites hacia nuevos horizontes. Con estos cambios tecnológicos los alumnos reclaman la presencia online del docente. En este sentido se podría pensar que contar con instrumentos automatizados e interactivos como es el caso de los programas robots conversacionales o *chatbots*. La inclusión de este tipo de tecnologías como soporte al proceso de enseñanza-aprendizaje resultará claramente beneficioso para el alumnado y profesorado.

Este proyecto plantea el diseño y construcción de *chatbots* educativos como un elemento de ayuda y soporte a la labor docente del profesorado universitario de la Facultad de Ingeniería de la Universidad Nacional de Jujuy. La motivación viene dada por la necesidad de contar con un mecanismo automatizado que permita satisfacer rápidamente los requerimientos de los alumnos relacionados con resolver dudas y consultas de una asignatura específica en el momento que lo precise.

Palabras clave: *ChatBot; Entorno virtual; Comunicación virtual; PLN; Interacción*

CONTEXTO

La línea de investigación de I+D+i aquí presentada nace como iniciativa del grupo de investigación VRAIn (Visualización y Recuperación Avanzada de Información) de la Facultad de Ingeniería de la Universidad Nacional de Jujuy (UNJu), cuyo eje principal es la mejora del proceso de enseñanza de asignaturas teóricas en carreras tecnológicas. Puesto que el equipo de trabajo forma parte del plantel docente de las tres carreras informáticas de la Facultad de Ingeniería: Ingeniería Informática, Licenciatura en Sistemas y Analista Programador Universitario; somos conscientes del beneficio que representará para los alumnos contar con estos avances.
Desde el punto de vista de los recursos humanos, este proyecto contribuirá a mejorar la formación en temas de Inteligencia Artificial, Aprendizaje Automático y Procesamiento del Lenguaje Natural.

1. INTRODUCCIÓN

La incursión de las Tecnologías de la Información y Comunicación (TIC) en la educación modificó las prácticas de enseñanza tradicionales, dando apertura a modalidades educativas no convencionales, a distancia, online, híbridas y a mecanismos de interacción tanto síncronos y asíncronos donde el docente y estudiante mantienen encuentros en entornos virtuales distintos al aula (Islas Torres, 2015). Este cambio posibilitó la aparición de conceptos como *b-learning* (*blended learning*, aprendizaje combinado), una modalidad de enseñanza-aprendizaje prometedora puesto que el híbrido de presencialidad y virtualidad incrementa las opciones para la construcción de conocimiento. En ese mismo sentido, como una aplicación de este nuevo concepto, se planteó un cambio de lugar entre las dos principales actividades del modelo tradicional: "los deberes en clase y las lecciones en casa", que pasó a denominarse Aula Invertida (Inverted Classroom, Classroom Flip, Flipped Classroom, Flipped Learning) (Baker, 2000; Lage, Platt, & Treglia, 2000; Lacleta, Luisa, Fidalgo Blanco, & García-Peñalvo, 2015).

El aspecto común de estas nuevas denominaciones es que las clases se sustituyen por videos, lecturas o cualquier otra actividad que el alumnado pueda hacer de forma independiente fuera del horario de clase y el aula se transforma en un punto de encuentro donde se realizan actividades prácticas y de forma cooperativa. Si bien no se puede afirmar que todas las clases de las carreras impartidas en la UNJu adoptan este nuevo concepto de forma total (aula invertida), es necesario destacar la importancia y relevancia del uso del Aula Virtual que al menos la Facultad de Ingeniería y sobre todo las carreras tecnológicas, han implementado desde hace años (desde el año 2011 con UnjuDigital[1] y desde el año 2015 con VirtualUnju[2], actualmente UnjuDigital está migrando a VirtualUnju como única plataforma oficial de educación a distancia).

El empleo del Aula Virtual funciona brindando soporte a la labor realizada por los docentes, principalmente como un repositorio de material educativo y como un entorno virtual que posibilita la comunicación y la interacción docente-alumno. El Aula Virtual aporta ciertas ventajas frente al modelo de enseñanza tradicional, por ejemplo, elimina las barreras espacio-temporales, ofrece agilidad y rapidez para solventar dudas, permite acceder a los tutores sin necesidad de desplazamientos, presenta un medio de trabajo colaborativo entre los alumnos y claramente actúa de facilitador para acceder a la información.

Si bien el Aula Virtual ofrece espacios para interactuar (foros de conversación por ejemplo), el alumno prefiere expresar su duda o consulta a través del mensaje privado, se ha detectado que al alumno le cuesta iniciar el debate/consulta/pregunta, principalmente por la timidez y por el miedo a quedar expuesto.

Por otro lado, a pesar de contar con estos entornos virtuales, el alumno reclama la presencia online del docente más allá del correo electrónico, los foros o mensajería instantánea. En este sentido se puede apelar a mecanismos más interactivos y eficientes que pueden operar con cierta autonomía con una amplia disponibilidad de acceso, como los programas robots conversacionales, o *chatbots* (Batista, 2017).

[1] http://www.unjudigital.unju.edu.ar/

[2] https://virtual.unju.edu.ar/

Los asistentes conversacionales o mejor conocidos como *bot* o *chatbot* o *chatterbot*, al que se puede definir como un asistente virtual, es un conjunto de programas informáticos que posee la habilidad de mantener una conversación con un ser humano mediante el lenguaje natural (Abdul-Kader & Woods, 2015). La inclusión de este tipo de tecnologías como soporte al proceso de enseñanza-aprendizaje resultará claramente beneficioso para el alumno.

El uso de *chatbots* dentro el ámbito escolar no es nuevo, han habido algunos acercamientos con resultados muy satisfactorios (Jia, 2009; Kerly, Ellis, & Bull, 2008; Kerly, Hall, & Bull, 2007; Hiremath, Hajare, Bhosale, Nanaware, & Wagh, 2018; Clarizia, Colace, Lombardi, Pascale, & Santaniello, 2018), y en cada uno de ellos se destaca el propósito de cada *chatbot*, es decir, cada *chatbot* es específico de un área de conocimiento o el menos si está dedicado a una sola área, los resultados mejoran notablemente, esto responde al hecho de que las técnicas empleadas para proporcionar el dialogo y sobre todo la respuesta a una pregunta dada, mejoran cuando se trata de un ámbito específico en contraposición a un *chatbot* de propósito genérico.

En una etapa inicial se planea realizar la construcción de un *chatbot* para una asignatura de la carrera Analista Programador Universitario, la asignatura en cuestión cuenta con alrededor de 80 alumnos y posee abundante material teórico. En este sentido se plantea automatizar las respuestas a dudas o consultas que el estudiante pueda tener en cualquier momento, puesto que estas herramientas se encuentran operativas las 24hs del día. La asignatura a evaluar posee mucho contenido teórico y conceptual, por tanto es una candidata ideal para este tipo de experimentos en dónde el Procesamiento del Lenguaje Natural (PLN) se realiza para hallar relaciones semánticas entre la pregunta realizada y la respuesta que eventualmente brindaría la herramienta.

La herramienta que se propone desarrollar se encuadra en el campo de la Inteligencia Artificial, específicamente lo que corresponde a PLN (Manning, 1999), puesto que un *chatbot* interactúa mediante el lenguaje natural, este tipo de técnicas resultan adecuadas para lograr una correcta aproximación.

La construcción del primer prototipo se realizará íntegramente en Python[3], empleando la librería *Gensim* (Rehurek & Sojka, 2010) que posee implementaciones de los algoritmos más comúnmente utilizados para PLN. En este caso se seleccionaron *Term Frequency - Inverse Document Frequency* (TF-IDF) (Aizawa, 2003), *Latent Semantic Indexing/Latent Semantic Analysis* (LSI/LSA) (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990; Landauer & Dumais, 1997) y *Word Mover's Distance* (WMD) (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013; Kusner, Sun, Kolkin, & Weinberger, 2015)

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Este trabajo tiene como principal línea de investigación el diseño, desarrollo, implementación y puesta en marcha de *chatbots* para el ámbito académico. Claramente es un trabajo interdisciplinario pues requiere de conocimiento en programación, desarrollo web, bases de datos, manejo de técnicas de

[3] https://www.python.org/

aprendizaje automático, procesamiento del lenguaje natural, entre otras.

Los continuos avances tecnológicos, la generación de información y la abundancia de dispositivos hiperconectados requieren la inclusión de herramientas e instrumentos como los que se proponen con este proyecto. Por esta razón, la inclusión de un *chatbot* dentro de una asignatura, representará un avance importante en el proceso de enseñanza-aprendizaje.

3. RESULTADOS OBTENIDOS/ESPERADOS

Tras el desarrollo del presente proyecto se espera:

- Poder evaluar el compromiso de los alumnos con el uso responsable de la herramienta.
- Analizar el comportamiento de los alumnos identificando temáticas, dudas o preguntas similares que permitan establecer ciertos patrones, a partir de la información almacenada al momento de utilizar la herramienta.
- Definir los lineamientos tecnológicos para la correcta implementación y despliegue de un *chatbot* para una asignatura.
- Establecer el formato de la información a ser utilizada y procesada por los modelos de predicción diseñados.
- Establecer el protocolo de reentrenamiento de los modelos de predicción diseñados.

4. FORMACIÓN DE RECURSOS HUMANOS

Dentro de la temática presentada, todos los miembros del proyecto imparten docencia en carreras informáticas, con lo cual, el desarrollo de este proyecto brinda un marco para que tanto los docentes como los estudiantes lleven a cabo tareas de investigación y se desarrollen en el ámbito académico.

En el transcurso del proyecto se tiene como objetivo consolidar la formación en investigación de los integrantes con menos antecedentes en proyectos y también está contemplado que los alumnos del último año de las carreras informáticas de grado, pueden optar por realizar su trabajo fin de grado con temáticas afines.

5. BIBLIOGRAFÍA

Abdul-Kader, S. A., & Woods, J. C. (2015). Survey on chatbot design techniques in speech conversation systems. *International Journal of Advanced Computer Science and Applications, 6*(7).

Aizawa, A. (2003). An information-theoretic perspective of TF-IDF measures. *Information Processing & Management, 39*(1), 45-63.

Baker, J. W. (2000). The 'Classroom Flip': Using Web Course Management Tools to Become the Guide by the Side en J. A. Chambers (Ed.). *11th International Conference on College Teaching and Learning*, (págs. 9-17). Jacksonville, FL.

Batista, A. (2017). Uso de chatbots como apoyo para la comunicación en el Aula. Un asistente virtual 24x7x365 colaborando con el curso. *4° Jornadas de TIC e Innovación en el Aula.* La Plata.

Clarizia, F., Colace, F., Lombardi, M., Pascale, F., & Santaniello, D. (2018). Chatbot: An Education Support System for Student. *Cyberspace Safety and Security*, (págs. 291-302).

Deerwester, S., Dumais, S., Furnas, G., Landauer, T., & Harshman, R. (1990). Indexing by latent semantic analysis. *Journal of the American society for information science, 41*(6), 391.

Hiremath, G., Hajare, A., Bhosale, P., Nanaware, R., & Wagh, K. S. (2018). Chatbot for education system. *International Journal of Advance Research, Ideas and Innovations in Technology, 4*(3), 37-43.

Islas Torres, C. (2015). La interacción en el b-learning como posibilitadora de ambientes de aprendizaje constructivistas: perspectiva de estudiantes. *Pixel-Bit. Revista de Medios y Educación*(47), 7-22.

Jia, J. (2009). CSIEC: A computer assisted English learning chatbot based on textual knowledge and reasoning. *Knowledge-Based Systems, 22*(4), 249-255.

Kerly, A., Ellis, R., & Bull, S. (2008). CALMsystem: a conversational agent for learner modelling. *Applications and Innovations in Intelligent Systems XV* (págs. 89-102). Springer.

Kerly, A., Hall, P., & Bull, S. (2007). Bringing chatbots into education: Towards natural language negotiation of open learner models. *Applications and Innovations in Intelligent Systems XIV*, (págs. 179-192).

Kusner, M., Sun, Y., Kolkin, N., & Weinberger, K. (2015). From word embeddings to document distances. *International Conference on Machine Learning*, (págs. 957-966).

Lacleta, S. E., Luisa, M., Fidalgo Blanco, Á., & García-Peñalvo, F. J. (2015). Metodología de enseñanza inversa apoyada en b-learning y gestión del conocimiento. *III Congreso Internacional sobre Aprendizaje, Innovación y Competitividad*, (págs. 464-468).

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment. *The Journal of Economic Education, 31*(1), 30-43.

Landauer, T., & Dumais, S. (1997). A solution to plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review, 104*(2), 211.

Manning, C. D. (1999). *Foundations of statistical natural language processing.* MIT press.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Neural Information Processing Systems*, 3111-3119.

Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks.*

Evaluación y Selección de Software Educativo utilizando el método LSP

Estela Fritz[1]; María Eva Ascheri[1]; Alejandra Zangara[2]

[1]Departamento de Matemática
Universidad Nacional de La Pampa
Av. Uruguay 151 – (6300) Santa Rosa – La Pampa – Argentina
Tel.: +54-2954-245220 – Int. 7125
[fritzem, mavacheri]@exactas.unlpam.edu.ar

[2] Facultad de Informática
Universidad Nacional de La Plata
50 y 120 – (1900) La Plata – Argentina
Tel.: +54-221-4277270 / 4277271
alejandra.zangara@gmail.com

Resumen

El presente trabajo se enmarca en el Proyecto de Investigación: *"Propuesta de Clasificación de software libre utilizado en la enseñanza de la programación"*, que se desarrolla en el ámbito de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa.

El principal objetivo de este trabajo es generar un solo valor como resultado de la evaluación de un software, haciendo uso del método *LSP*. Este trabajo presenta además, una propuesta para incluir algunos criterios pedagógicos relacionados con la evaluación y selección del software educativo.

Palabras clave: Método LSP, evaluación de software, software educativo, métodos cuantitativos.

Contexto

Por Resolución N° 160/18 del Consejo Directivo de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa se acredita el Proyecto de Investigación: *"Propuesta de Clasificación de software libre utilizado en la enseñanza de la programación"*. El mismo es dirigido por la Magister María Eva Ascheri de la Facultad de Ciencias Exactas y Naturales de la Universidad Nacional de La Pampa (U.N.L.Pam) y co-dirigido por la Magister Alejandra Zangara de la Facultad de Informática de la Universidad Nacional de La Plata (UNLP). El mencionado proyecto incluye a la Profesora Estela Marisa Fritz como investigadora. Constituye su Trabajo Final para alcanzar el grado de Especialista en Tecnología Informática Aplicada en Educación. Además, será la base para desarrollar su futura tesis de maestría a llevarse a cabo en la UNLP.

1. Introducción

Los rápidos avances en Tecnología de la Información y la Comunicación sumados a la diversidad de software educativo, tornan casi imposible tomar decisiones que conduzcan a una mejora académica óptima. Los criterios para la selección del software educativo deben ser considerados en una instancia previa al proceso de evaluación y selección del software. Identificar apropiadamente esos criterios es la clave para

la evaluación exitosa y posterior selección del software.

Desafortunadamente, no hay una lista genérica de criterios que puedan ser usados para evaluar cualquier software [1]. Los investigadores generalmente proveen una lista de criterios relacionados con la correctitud, eficiencia, usabilidad [2] y proponen métodos de evaluación basados en encuestas al usuario, cuyas respuestas organizan características según una jerarquía: buena, muy buena, aceptable, no tan buena, entre otras.

Sin embargo, cuando evaluamos y seleccionamos software educativo, hay otros aspectos relacionados con el proceso de aprendizaje, que deben ser considerados, a saber:

1. El rol desempeñado por el docente en el contexto de uso del software.
2. Contribución del software educativo al desarrollo de habilidades intelectuales, estrategias cognitivas, información verbal, actitudes y otras.
3. La aproximación a los principios de enseñanza y aprendizaje subyacentes en el sistema de software.

Es por eso que los criterios técnicos ([3], [4]) son insuficientes para determinar la aptitud del software, y los criterios educacionales también deben ser considerados ([2], [5]). Ambas clases de criterios están relacionadas.

Como fue mencionado antes, los criterios de selección de software educativo incluyen tanto aspectos técnicos del software como pedagógicos.

1.1 Algunos aspectos del diseño instruccional

Antes de hacer una reseña del método LSP, se consideran algunos tópicos sobre "Principios de diseño instruccional" [6]. Se mencionan aquellos considerados relevantes para establecer criterios elementales y complejos para LSP. En principio, la identificación de los objetivos o metas instruccionales hasta el proceso de evaluación de estos. Se mencionan también las cinco categorías principales de resultados del aprendizaje, tales como habilidades intelectuales, estrategias cognitivas, información verbal, actitudes y habilidades motoras. La medida en la que el software objeto de la evaluación contribuya a desarrollar alguna o algunas de estas categorías de resultados del aprendizaje, es una característica para considerar como parte de la evaluación. Una mención especial merecen las condiciones de aprendizaje necesarias para la adquisición de esas capacidades. Las diferencias individuales entre los sujetos que aprenden y cómo esas diferencias afectan a la planificación de los escenarios de aprendizaje, es un tópico para incluir entre los criterios considerados cuando se evalúa un software.

También deben considerarse criterios relativos a las posibilidades de un software de incluir procedimientos de evaluación.

El aprendizaje es un proceso y como tal, provoca cambios en las capacidades del sujeto que aprende. Por eso, es deseable que un software educativo promueva dichos cambios.

Cuando se examina cuidadosamente, la situación de aprendizaje tiene dos partes: una externa al sujeto y una interna al sujeto que aprende. Esa parte interna deriva de aquello que el sujeto puede recordar o recuperar de su memoria. Esas partes (la externa y la interna) están relacionadas. Esa relación existente se denomina *condiciones del aprendizaje*.

1.2 Algunos principios del aprendizaje

Algunos de estos principios expanden las ideas expuestas. Uno de ellos es el conocimiento de las condiciones del aprendizaje humano. Aún hay tres principios derivados de la teoría del aprendizaje, que son relevantes: (1) Contigüidad: las situaciones de estímulo deben ser presentadas en el mismo tiempo que las respuestas esperadas. (2) Repetición: esa situación de estímulo y la respuesta del sujeto, necesitan ser repetidas. (3) Refuerzo: para los propósitos de la

instrucción, es necesario que un nuevo aprendizaje sea seguido inmediatamente por otro ya logrado y que el sujeto que aprende sea capaz de relacionarlos, de modo que aquel aprendizaje ya logrado se constituya en un refuerzo de lo nuevo.

Estos tres principios deberían ser incluidos en el proceso de evaluación y selección de un software educativo.

1.3 Reseña del método *LSP*

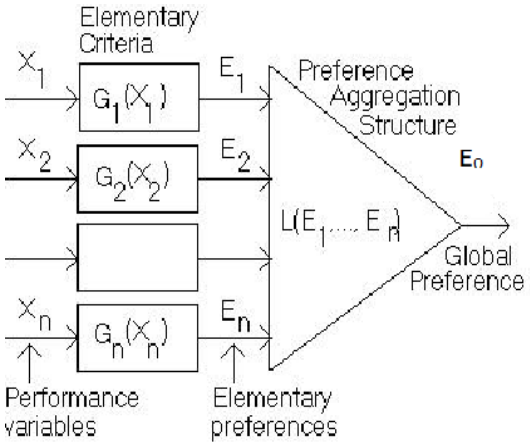
El método LSP (Logic Scoring Preferences) [7] es una generalización y extensión de varias técnicas de scoring. Es un método cuantitativo basado en el empleo de la lógica continua que permite la creación de funciones complejas de evaluación y su aplicación en la evaluación, optimización, comparación y selección de sistemas de software de propósito general [8].

Una vez identificadas las características que intervienen en la evaluación del software, los datos obtenidos de las distintas características del software, de acuerdo con *LSP*, serían las variables de performance del sistema. En el caso del software educativo, algunas variables de performance que se pueden describir son las siguientes: en qué medida es tenido en cuenta el aprendizaje individual, cuantificar fases de aprendizaje inmediato y de largo alcance, gradualidad del software para el aprendizaje (¿Es posible identificar esa gradualidad?), en qué medida el software está basado en el conocimiento de cómo aprende el ser humano, en qué grado contribuye el software a los principios de contigüidad, repetición y refuerzo, qué contenidos de la memoria o capacidades aprendidas el software contribuye a recuperar, qué categorías (en cuanto a resultados del aprendizaje) el software tiende a desarrollar, cuál es el rol del docente en el proceso, en qué medida el sistema guía al sujeto hacia el autoaprendizaje. Esto por mencionar algunas.

Los valores de dichas variables son mapeados por medio de funciones denominadas *criterios elementales*, en *preferencias elementales*. Son estas preferencias las que van siendo agregadas mediante operadores de la Lógica continua en estructuras de agregación que nos permiten obtener un único valor final (*preferencia global final* E_0 en la fig. 1).

Los criterios elementales son funciones que transforman un valor real de una variable de performance, en un valor perteneciente al intervalo [0,100] o [0,1] y expresan el grado de cumplimiento con un requisito del sistema que está siendo evaluado.

Figura 1. El proceso de evaluación de LSP



2. Líneas de Investigación y Desarrollo

El proyecto de Investigación citado aborda los siguientes aspectos relacionados:

- Relevamiento de software libre utilizado en la enseñanza de la programación.
- Descripción de un método cuantitativo para la evaluación y selección del software basándose en criterios tanto técnicos como pedagógicos.
- Evaluación del software relevado con el fin de clasificarlo según las habilidades intelectuales y estrategias cognitivas que permite desarrollar en el sujeto que aprende.

El presente trabajo corresponde al segundo de los aspectos mencionados.

3. Resultados Obtenidos/Esperados

Después de considerar los tópicos explicados anteriormente, como resultado, puede decirse que el método *LSP* propuesto constituye un muy buen método cuantitativo que incrementa la eficiencia de la evaluación y la selección de sistemas de software en general y software educativo en particular. Representa un método más ventajoso cuando se lo compara con otros métodos cuantitativos [10]. *LSP* tiene principalmente las siguientes ventajas relativas a los métodos cuantitativos para evaluación de sistemas:
(1) La especificación de requerimientos es sistemática, flexible y completa.
(2) El proceso de evaluación sistemático y racional, el cual refleja explícita y cuantitativamente el nivel global de satisfacción de los requerimientos del usuario.

El método *LSP* puede ser usado para la evaluación y comparación de una amplia variedad de sistemas complejos [11]. El principal poder de la estrategia de *LSP* es la capacidad de construir un modelo versátil de un conjunto de preferencias. Mediante la combinación apropiada de operadores (lógicos) de agregación con el conjunto de preferencias, es posible derivar criterios complejos que tienen poder expresivo y flexibilidad.

El software educativo debe satisfacer una variedad de requerimientos. El nivel global de satisfacción de esos requerimientos (**E_0**) es usado para su evaluación y comparación.

Por otro lado, se espera que esta investigación brinde a los docentes herramientas cuantitativas para la evaluación y selección de software, en función de los objetivos que desean alcanzar.

4. Formación de Recursos Humanos

En el marco del presente proyecto, como ya se mencionó anteriormente, se espera que la Prof. Fritz alcance el grado de Especialista en Tecnología Informática Aplicada en Educación y que, en un futuro, pueda desarrollar su tesis de maestría en la UNLP.

Actualmente, la Prof. Fritz ha completado la etapa de búsqueda y análisis de métodos de evaluación del software, tema en el que no era experta. Durante 2019, realizará todos los aportes relacionados con la definición de criterios para evaluar software educativo para la enseñanza de la programación, principalmente desde el punto de vista pedagógico, y tomando como base alguna teoría del aprendizaje en particular. Esto constituye un insumo necesario para los nuevos avances en el proyecto.

5. Bibliografía

[1] **Abohamad, W., Arisha, A.** (2010) "Evaluating and Selecting Optimization Software Packages: A Framework for Business Application" World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering. Vol. 4 https://waset.org/publications/4426/evaluating-and-selecting-optimization-software-packages-a-framework-for-business-applications

[2] **Gorga, G., Madoz, M., Pesado P**, (2000) "Hacia una propuesta de métrica para la evaluación de Software Educativo", CACIC.
En línea: http://hdl.handle.net/10915/23514

[3] **International Standard. ISO/IEC 25012** (2008) Software Engineering – Software Quality Requirements and Evaluation (SQuaRE) - Data Quality Model.

[4] **International Standard ISO/IEC 25010** (2011) Systems and software engineering — Systems and software Quality

Requirements and Evaluation (SQuaRE) — System and software quality models.

[5] **Cataldi, Z.,** (2000) "Una Metodología para el Diseño, Desarrollo y Evaluación de Software Educativo", Universidad Nacional de La Plata, Facultad de Informática, Tesis de Maestría.
En línea: http://hdl.handle.net/10915/4055

[6] **Gagné, R.M., Briggs, L., Wager, W**. (1992) Ed. Harcourt Brace College Publishers, 4th edition.

[7] **Dujmovic, J.J and Elnicki, R** (1982) "A DMS Cost/Benefit Decision Model: Mathematical Models for Data Management System Evaluation, Comparison, and Selection" National Bureau of Standards, Washington. DC., No NBS-GCR, NTIS No PB82- 170150 (155 pages).

[8] **Dujmovic, J.J** "A Method for Evaluation and Selection of Complex Hardware and SoftwareSystems
https://telin.ugent.be/~gdetre/Summer%20School%20Poland/Resources/Decision%20support.pdf

[9] **Daso, A**. **et al**, (2013) "Desarrollo de Modelos de Evaluación Usando Operadores de una Lógica Continua" XV WICC (Workshop de Investigadores en Ciencias de la Computación), págs. 420-424.

[10] **Squires,D., McDougall**, A.(1997) **"**Cómo elegir y utilizar software educativo", Ediciones Morata y Fundación Paideia, Primera Edición.

[11] **Dujmović, Jozo J., and Hajime Nagashima**. (2006) "LSP Method and Its Use for Evaluation of Java IDEs." International Journal of Approximate Reasoning 41.1: 3–22. Web.

EXPERIENCIA DE UN MOOC. El desafío de contar con aspectos de usabilidad y accesibilidad de la plataforma

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 desde hace varios años en el LINTI, Laboratorio de Investigación en Nuevas Tecnologías Informática de la Facultad de Informática de la UNLP, en relación a las plataformas de e-learning y al concepto de aprendizaje basado en recursos abiertos.

En esta etapa se está trabajando con MOOCs, fenómeno que surge en un contexto amplio de educación abierta y que comenzó a impactar en la educación superior en épocas de globalización. Continuando con lo expuesto en WICC 2018, y a partir de la experiencia del MOOC creado sobre Accesibilidad Web, se están analizando los resultados de dos ediciones del curso, respecto a la usabilidad de la plataforma y a la naturaleza y organización del contenido y de las evaluaciones incluidas. Por otro lado, se evaluarán las pautas de accesibilidad de la herramienta utilizada y en particular del curso implementado, con el fin de realizar las modificaciones necesarias.

Palabras clave: Recursos educativos abiertos, OER, MOOC, accesibilidad

CONTEXTO

La Facultad de Informática tiene una amplia experiencia en el uso de sistemas de gestión de aprendizaje, en especial la plataforma de código abierto Moodle. Esta plataforma es la base 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 abiertos y masivos, es que se comenzó a utilizar la plataforma edX para el desarrollo de MOOCs. Como primera experiencia se implementó 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 años son “Trabajando por una Web Accesible” [1] y “Trabajando por una Web Inclusiva: un desafío que nos compromete a todos” [2].

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.

1. INTRODUCCIÓN

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. Si bien pueden ser considerados como una extensión de los enfoques de aprendizaje electrónico existente, los MOOCs ofrecen la oportunidad de repensar acerca de los nuevos modelos de educación superior en línea y gratuita, en donde los cursos están abiertos a cualquier usuario de internet [4]

Los MOOCs surgieron como consecuencia de las nuevas tendencias internacionales en el aprendizaje basado en la tecnología, como parte del movimiento educativo abierto [5]. 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 [6].

El diseño de un MOOC es un proceso complejo 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 [7]. En muchos casos, hay suficientes materiales y ejercicios a disposición de los alumnos desde que ingresan en el curso, que les permiten gestionar su aprendizaje según su propio ritmo. Son muy importantes los aspectos motivacionales para evitar el abandono de los cursos y favorecer el aprendizaje. 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.

Por otro lado, los MOOCs imponen 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. Gran parte del contenido suele ser multimedia basado en video y en muchos casos incluyen juegos serios.

Respecto a la Accesibilidad, es una temática compleja que reúne características de índole social, moral, legal, educativa como técnica. Va desde los derechos de las personas con discapacidad al acceso a la información, normas y recomendaciones internacionales sobre accesibilidad web, como cuestiones de implementación en el código. Hay un consenso general que es necesario analizar y plantear políticas de accesibilidad en plataformas que albergan Recursos Educativos Abiertos (Open Educational Resources-OER), lo que significa que los repositorios abiertos deben ser diseñados teniendo en mente el concepto de accesibilidad [8]. Los MOOCs están estrechamente relacionados con los OER en tanto

representan contenido abierto al que se le agrega la característica de masividad.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Respectos a las tareas de investigación, desarrollo e innovación que se vienen llevando a cabo en este proyecto se encuentran distintas líneas de trabajo relacionadas a e-learning, cuestiones de interoperabilidad e integración de plataformas de código abierto utilizadas en este campo y sus aspectos de accesibilidad. Se compararon distintas herramientas open source para soporte de MOOCs como OpenEdx [9] y openMOOC [10] 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. Usando esta plataforma se implementó un MOOC sobre Accesibilidad Web, que se puso en producción en septiembre de 2017, y se encuentra actualmente disponible en https://actividades.linti.unlp.edu.ar. Sobre la plataforma en cuestión 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 [11].

Continuando con las líneas planteadas en WICC 2018 [12], en esta etapa el foco de estudio se relaciona con la evaluación del caso de uso desarrollado a partir de las experiencias llevadas a cabo durante los años 2017 y 2018. La evaluación se basa en analizar el impacto del contenido y de las actividades y ejercitaciones planteadas a través de las experiencias y devoluciones de los alumnos participantes. El objetivo es establecer un punto inicial en la tarea de diseñar y crear cursos masivos de calidad.

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 [13]. Se comenzó realizando un análisis global del curso y de la plataforma utilizada, en forma interrelacionada, en base a diferentes metodologías de diseño planteadas en otras instituciones [7] [14]. Respecto del curso se consideran características propias de contenido y lo relacionado a la secuencia de aprendizaje. En este punto, se analiza en detalle, tanto la organización del curso en las unidades planteadas, como el formato o tipo de contenido incluido en cada una de ellas. Según lo analizado, en relación al contenido, se incorporará más contenido multimedial y se reestructurará su organización, planteando diferentes secuencias de acuerdo al perfil del alumno. Esto le permitirá personalizar el trayecto de aprendizaje según sus necesidades. Se analizan 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. 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 analizan 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. Las evaluaciones, que se continúan realizando, se basan en guías metodológicas utilizadas en universidades del exterior y en distintos estudios realizados [15] [16].

Respecto a las características de la plataforma, se están evaluando 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. Esto requiere aplicar diferentes estrategias pedagógicas para poder, en el proceso de enseñanza a distancia, concientizar a los estudiantes, proponer el ponerse en el lugar del otro, transmitir empatía y valores éticos, incentivar la investigación de la problemática de los usuarios con discapacidad y trabajar por y para garantizar una experiencia del usuario inclusiva, promover la colaboración para la búsqueda conjunta de soluciones accesibles y formas de aplicarlas en las producciones informáticas que se desarrollen e instar a la evaluación y comprobación continua de la accesibilidad. Además, en la evaluación de los aspectos de accesibilidad se considera tanto las expectativas de los alumnos como el punto de vista y la experiencia de los docentes.

La implementación del curso nos sirve como experiencia para transformar en MOOCs diversos cursos en línea existentes en un LMS tradicional. Este proceso involucra analizar y evaluar también nuevas plataformas de base, nuevos medios de comunicación, contenidos y formatos. En esta línea se realiza un trabajo interdisciplinario con diseñadores visuales, expertos en comunicación, contenidos y educadores que permite 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 descritas, se plantean los siguientes objetivos:

- Estudiar distintas metodologías de diseño de MOOCs aplicadas en universidades del mundo, analizando sus ventajas y posibles adecuaciones al ámbito local.
- Evaluar el MOOC creado desde distintas perspectivas docente, institucional y técnica, a partir del estudio de distintas experiencias, analizando las opiniones de los alumnos involucrados, fomentando el aprendizaje centrado en el estudiante.
- Como base de la planificación estipulada, analizar los resultados obtenidos en cada etapa, corrigiendo posibles falencias en un ciclo de mejora continua
- Utilizar técnicas de diseño estudiadas en futuros cursos, a partir del caso de uso desarrollado.
- Llevar a cabo estudios comparativos entre la experiencia del MOOC y los cursos en un LMS tradicional, que permitan

aportar nuevos conceptos a los entornos de aprendizaje y enseñanza que se generan con estas herramientas.

- Redefinir componentes del proceso de aprendizaje, relacionadas a la naturaleza del contenido y al diseño temporal, a fin de adaptar el curso según las evaluaciones realizadas.
- 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.
- Realizar un análisis continuo de las pautas de accesibilidad, a través de evaluaciones heurísticas utilizando herramientas específicas y empíricas teniendo en cuenta el perfil del usuario, ya sea estudiante, docente, tutor o administrador. Evaluar la posibilidad de intervención y participación de una persona con discapacidad en los diferentes roles.
- Establecer pautas de diseño y construcción de MOOCs que se utilicen como punto de partida para la la creación de cursos masivos 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] González, C., Collazos, C., García, R. “Desafío en el diseño de MOOCs: incorporación de aspectos para la colaboración y la gamificación”. RED - revista de Educación a Distancia. Núm. 48. Art. 7. 30-Ene-2016. DOI: 10.6018/red/48/7

[5] Ramírez, M., Burgos, J. “Movimiento Educativo Abierto. Acceso, colaboración y movilización de recursos educativos abiertos.” México: CIITE., 2012

[6] 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.

[7] 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).

[8] Law, P., Perryman, L.-A., and Law, A., 2013. Open educational resources for all? Comparing user motivations and characteristics across The Open University's iTunes U channel and OpenLearn platform.

[9] Open edX Portal | Open Source MOOC Platform https://open.edx.org

[10] Open MOOC https://www.openeducationeuropa.eu/es/institution/open-mooc

[11] Accessibility Features | Open edX Portal | Open Source MOOC Platform https://open.edx.org/features-roadmap/accessibility/all

[12] Díaz, J., Schiavoni, A., Amadeo, P., Harari, I., Osorio, A., Carrilao Ávila, F. "Análisis de aspectos de diseño en cursos masivos en línea. Caso de estudio sobre un curso de Accesibilidad Web". WICC 2018, XX Workshop de Investigadores en Ciencias de la Computación, ISBN: 978-987-3619-27-4, Corrientes, Argentina, 26-27 Abril, 2018.

[13] 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

[14] "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.

[15] 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

[16] Daradoumis, T., Bassi, R., Xhafa, F., Caballé, S. "A review on massive e-learning (MOOC) design, delivery and assessment". 2013 Eighth International Conference on P2P, Parallel, Grid, Cloud and Internet Computing, France, 2013. IEE Xplore. Electronic ISBN: 978-0-7695-5094-7

Framework para la generación de Video juegos educativos en Sistemas de Realidad Aumentada- Líneas de Investigación del Grupo de Realidad Aumentada Aplicada

Martin Becerra, Jorge Ierache, Nahuel Mangiarua, Hernan Maurice, Santiago Igarza, Osvaldo Spositto

Universidad Nacional de La Matanza, Departamento de Ingeniería e Investigaciones Tecnológicas.

Grupo de Realidad Aumentada Aplicada.

Florencio Varela 1903, La Matanza, Buenos Aires, Argentina
jierache@unlam.edu.ar

Resumen

El presente proyecto se orienta al desarrollo de juegos educativos en el contexto de realidad aumentada a fin de generar juegos educativos con una mecánica similar al de juego la oca, alimentado por los contenidos publicados en el sistema de catálogos aumentados, facilitando una herramienta lúdica dinámica y colaborativa para entornos educativos. Se propone que el editor web del sistema de catalogo aumentado permita al usuario armar categorías de preguntas y respuestas a partir de contenidos virtuales y marcadores del sistema para que sean utilizados en el visor de realidad aumentada. El mismo se busca ampliarlo con una nueva funcionalidad que permita reproducir la mecánica del juego de preguntas y respuestas.

Palabras clave: Realidad Aumentada, Video Juegos, Educación.

Contexto

La investigación 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-213.

El grupo se encuentra financiado a través del desarrollo de un módulo del sistema de catálogos virtuales aumentados para generar video juegos educativos aumentados para el proyecto PROINCE 213.

Introducción

La realidad aumentada mediante teléfonos móviles es una de las subáreas de investigación de realidad aumentada que más está creciendo en la actualidad. Se define a un sistema de Realidad Aumentada (RA) como un sistema que permite a usuarios interactuar en tiempo real con contenidos virtuales asociados a objetos del entorno real [1],[2]. Podemos ampliar esta definición como un sistema que aumenta los sentidos de la persona (Visión, audio y tacto) para permitir la interacción con contenidos virtuales del mundo digital que son invisibles en el

entorno real [3]. El grupo de realidad aumentada aplicada realizo diversos trabajos en el campo de realidad aumentada que permite a usuarios sin mayores conocimientos técnicos de RA poder crear catálogos virtuales aumentados agrupando contenidos virtuales asociados a marcadores de RA como imágenes impresas[4],[5]. En el área de educación el grupo de investigación ha realizado diversas aplicaciones experimentales [6],[7] orientadas a juegos, como así también aplicaciones experimentales destinadas a la explotación de materiales didácticos para el área educativa [8],[9].

Se presenta en las secciones siguientes la línea de investigación Generación de video juegos educativos a partir de un catálogo creado en el sistema de catálogos virtuales aumentados, sus resultados y objetivos.

Líneas de Investigación, Desarrollo e Innovación

Generación de video juegos educativos

La línea de investigación se orienta al desarrollo de un módulo adicional para el Sistema de Catálogos Virtuales Aumentados [4],[5] extendiendo así el alcance de las funcionalidades de este. Se propone entonces la incorporación de generación de juegos educativos a partir de los catálogos de realidad aumentada creados en el sistema.

La nueva extensión permitirá al alumno, fijar los conceptos vistos en el aula mediante la creación de preguntas multiple choice referidas a temáticas de la materia. Las preguntas estarán agrupadas en diferentes temáticas asociadas a marcadores de RA, creados por el usuario generador de contenidos del catálogo virtual aumentado, con el fin de ser explotados en el juego de realidad aumentada.

La arquitectura basal del juego se desarrolla sobre el esquema del juego de la oca (Fig 1 tablero aumentado) donde los usuarios obtienen la cantidad de casilleros a avanzar mediante un dado virtual (Fig 2 dado y posición de los jugadores aumentada) y contestan preguntas aumentadas en el tablero para seguir avanzando. Tanto el uso del catálogo, como la nueva funcionalidad de creación de juegos educativos demuestran la necesidad de incorporar otras funcionalidades para facilitar la carga. Por este motivo, se explorará la posibilidad de integrar y ampliar el uso de templates desarrollado en [10] para poder facilitar la creación de categorías y contenidos para el mismo.

Figura 1 Tablero Aumentado

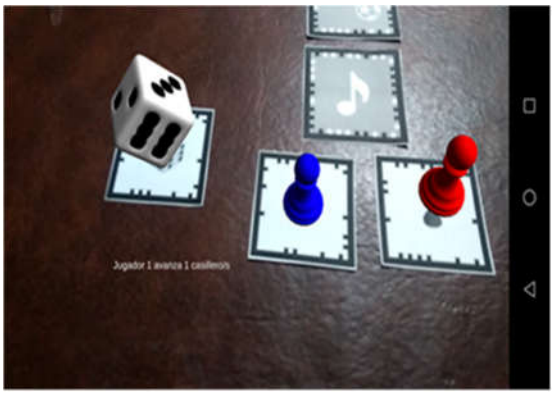


Figura 2 Dado y fichas aumentadas de los jugadores

Resultados y Objetivos

El objetivo general de esta investigación aplicada es contribuir con el desarrollo de una herramienta lúdica dinámica y colaborativa para entornos educativos. Por un lado, se buscará utilizar el editor web del sistema de catálogo aumentado para permitir al usuario armar categorías de preguntas y respuestas a partir de contenidos virtuales y marcadores del sistema para que sean consumidos utilizando un visor de realidad aumentada. Por otro lado, se busca ampliar las funcionalidades básicas existentes del visor de realidad aumentada para que reproduzca el juego de preguntas y respuestas.

Se plantearon como objetivos específicos: desarrollar el módulo de creación de catálogo de video juego para la plataforma web del Sistema de Catálogos Virtuales Aumentados y desarrollar la funcionalidad de la plataforma móvil del Sistema de Catálogos Virtuales Aumentados para explotar los diversos contenidos de categorías y preguntas creadas por los usuarios finales.

Los resultados hasta el momento permitieron realizar pruebas de diferentes implementaciones de la mecánica general del juego para encontrar un mecanismo que facilite a los usuarios implementaciones de la mecánica general del juego, estos fueron publicados recientemente [11]. Para encontrar un mecanismo que facilite a los usuarios poder utilizar el sistema las próximas actividades serán dedicadas a la integración del framework de generación de templates [10] a fin de facilitar la creación de categorías y contenidos para el mismo, como así también realizar pruebas con un grupo de usuarios que no posean conocimientos de realidad aumentada.

Formación de Recursos Humanos

El grupo de investigación se encuentra conformado por tres investigadores formados, tres investigadores en formación, y un alumno del departamento de Ingeniería e Investigaciones Tecnológicas de la UNLaM, trabajando en el área de Realidad Aumentada. Dos investigadores en formación se encuentran realizando el doctorado en ciencias informáticas en la UNLP:

- Nahuel A. Mangiarua: “Integración escalable de Realidad Aumentada basada en imágenes y rostros”. Directores: Jorge S. Ierache (UNLaM) y María José Abásolo (UNLP).

- Martin Becerra “Aumentación de Sistemas SCADA en el Contexto de la Industria 4.0”. Directores: Jorge Ierache (UNLaM) y María José Abásolo (UNLP).

Referencias

1. Papagiannakis G., Gurminder S. y Nadia M. T., “A survey of mobile and wireless technologies for augmented reality

systems." Comput. Animat. Virtual Worlds, vol. 19, no. 1, pp. 3-22, 2008.

2. Azuma R. T., The Most Important Challenge Facing Augmented Reality, Presence Teleoperators Virtual Environ., vol. 25, n.o 3, pp. 234-238, dic. 2016.

3. Specht M., Ternier S. y Greller W., "Dimensions of mobile augmented reality for learning: a first inventory,"Journal of the Research for Educational Technology (RCET), vol. 7, no. 1, pp. 117-127, 2011

4. Ierache J., Mangiarua N., Verdicchio N., Becerra M., Duarte N., Igarza S. "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. 2014

5. Ierache J., Mangiarua N., Bevacqua S., Verdicchio N., Becerra M., Sanz D., Sena M., Ortiz F., Duarte N., Igarza S. "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. 2015

6. Becerra M., Sanz D., Igarza I, Mangiarua N., Ierache J. "Sistema de Catálogo Virtual Aumentado. Integración de Framework Especializado orientado a Material Didáctico" TEYET, pp 350-356, ISBN 978-950-656-154-3. 2015

7. Mangiarua N., Ierache J., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. "Herramienta de Realidad Aumentada para la explotación de material didáctico tradicional". TE&ET. E-Book. ISBN 978-987-24611-1. 2014

8. Verdicchio N., Sanz D., Igarza S., Mangiarua N., Montalvo C., Ierache J. "Sistema de Catalogo Virtual Aumentado Integración de Framework Especializado Orientado a Juegos Didácticos". TE&ET, pp 597-604,ISBN 978-987-3977-30-5. 2016

9. Ierache J., Mangiarua N., Bevacqua S., Becerra M., Verdicchio N., Duarte N., Sanz D., Igarza S. "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. 2014

10. Mangiarua N., Montalvo C., Petrolo F., Sanz D., Verdicchio N., Lobatto E., Rosenthal A., Becerra M., Igarza S., Ierache J. "Framework para la Generación de Templates en Sistemas de Catálogos de Realidad Aumentada" XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017, Buenos Aires, Argentina) p. 393-397 ISBN:978-987-42-5143-5. 2017

11. Ierache J., Mangiarua N., Becerra M., Igarza S. Framework for the Development of Augmented Reality Applications Applied to Education Games. In: De Paolis L., Bourdot P. (eds) Augmented Reality, Virtual Reality, and Computer Graphics. AVR 2018. Lecture Notes in Computer Science, vol 10850. Springer, Cham. p. 340-350. 2018

Hacia el desarrollo de un sistema de Realidad Virtual
Estudio de caso: Centro de rehabilitación Integral de la provincia de San Juan

Mónica González 2, Cintia Ferrarini 2 ,Alicia Aballay2, 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, lolguin}@iinfo.unsj.edu.ar, [2] {gonzalez.monicagilda, ferrarinicintia,prof.alicia }@gmail.com

1. RESUMEN

El proyecto "*Videojuegos en realidad virtual como intervención terapéutica alternativa para estimular la cognición en niños con encefalopatía crónica no evolutiva*" (Vivitan), tiene como objetivo esencial desarrollar videojuegos, en ambientes de realidad virtual, para la estimulación de la atención y la memoria de niños y adolescentes con discapacidad cognitiva que asisten al Centro de Rehabilitación Integral (ubicado en el departamento de Rawson, provincia de San Juan), como terapia de intervención alternativa. Se pretende que el videojuego ofrezca un entorno inmersivo de aprendizaje basado en el descubrimiento y en la creatividad, proponiendo actividades lúdicas motivantes para estimular las funciones cognitivas: atención y memoria en niños y jóvenes con discapacidad cognitiva.

Actualmente en el Centro de Rehabilitación Integral las dificultades cognitivas son tratadas mediante intervenciones terapéuticas tradicionales. La motivación que despierta este tipo de videojuegos en sus usuarios, se considera que puede estimular la cognición y favorecer la psicomotricidad de niños con discapacidad. Al respecto, existen estudios científicos que demuestran que el uso de la realidad virtual podría contribuir en forma efectiva a dichas intervenciones terapéuticas.

Vivitan consiste en una investigación aplicada, en la que se ha optado por un diseño descriptivo y experimental, en el marco de una metodología cuanti-cualitativa. En el desarrollo del software se está utilizando una metodología ágil basada en Scrum. Las instancias de evaluación de usabilidad, jugabilidad y experiencia de usuario se realizarán con la participación de terapeutas y pacientes que asisten a la institución demandante.

En este trabajo se presenta el estado de avance del proyecto.

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 proyecto "*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 está llevando 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

Los videojuegos de Realidad Virtual proponen ambientes sensoriales enriquecidos para sus jugadores. Vega Muriel et. al. (2014) indica que existen numerosos estudios cuyos resultados demuestran que la exposición a estos ambientes mejora la función cognitiva en niños con ECNE. La inclusión de programas informáticos en tratamientos a niños con ésta patología, han mostrado una mejora en el funcionamiento espacial respecto a un grupo que no utilizó esta tecnología. Esto se debería a que la utilización de recursos tecnológicos genera una mayor motivación y por tanto, mejora la eficacia del tratamiento.

La realidad virtual (VR) propone un escenario inmersivo generado por computadora que simula una experiencia del mundo real, basada en la realidad o la ciencia ficción.

En (Howcroft. et al, 2012) se exponen resultados de una investigación en la que se demuestra que el uso de tecnología de realidad virtual (VR) aumenta la motivación de los niños y promueve la participación en las actividades propuestas, en comparación con los tratamientos convencionales de papel y lápiz.

González Sánchez, J. (2010) señala: "Cuando desarrollamos software para videojuegos, es crucial tener en cuenta que el usuario, a la hora de utilizarlo, sienta las mejores sensaciones, es decir que la experiencia de uso sea exitosa pues es el usuario el verdadero protagonista a la hora de interactuar con la aplicación."

Al respecto afirma Dierssen (2014):" El desafío, al trabajar con niños que sufren patologías que dificultan su aprendizaje, es desarrollar un videojuego con estética moderna, accesible sustentada en una base científica validada sobre la estimulación cognitiva".

Desde la perspectiva de Howcroft. et al, (2012), en el ámbito de la salud, existen estudios con resultados aceptables respecto a la utilización de VR en el tratamiento de patologías discapacitantes.

El gran desafío tecnológico de la VR es lograr el objetivo de promover el desarrollo del niño en un entorno divertido. En relación al juego Vigotsky afirma: "*Las funciones mentales superiores se adquieren y se desarrollan a través de la interacción social y el juego es un gran agente socializador*".

La naturaleza lúdica de los videojuegos en VR, estimulan la cognición favoreciendo el proceso de aprendizaje, dado que incrementan la atención y memoria a en la resolución de un problema concreto. (Howcroft J, Klejman S, Fehlings D, Wright V, Zabjek K, Andrysek J, 2012)

La atención, en términos generales, es un proceso multidimensional que permite seleccionar y concentrarse en una o varias tareas relevantes. En ese sentido, se toma como referencia para este proyecto el modelo conceptual sobre la atención propuesto por Sohlberg y Mateer (1987).

Por otra parte, la memoria es la capacidad de codificar, almacenar y recuperar de manera efectiva información aprendida o un suceso vivido. Dependiendo del tiempo que permanece la información en el sistema cognitivo se habla de memoria a corto plazo, de trabajo y de largo plazo. Por otro lado, dependiendo del tipo de información que la memoria retiene se habla de memoria verbal (lo que se escucha o se dice o lo que se lee en términos de palabras) o no verbal (lo que se ve, se toca, etc.). En relación con el tipo de órgano sensorial empleado se clasifica en memoria visual, auditiva (o ecoica), olfativa, gustativa, o háptica.

ViVITAN está orientado a desarrollar un sistema de juego de realidad virtual inclusivo, requiere la definición de un perfil de jugador con discapacidad cognitiva a fin de poder identificar y generar las interfaces adaptadas a sus necesidades. Además, pretende reducir las barreras en el acceso a los videojuegos para a las personas con discapacidad cognitiva.

Esto es posible si se consideran en el diseño del software la Cultura organizacional de la Institución demandante. Entendiendo por Cultura de la organización a aquella que comprende el "conjunto común de creencias, valores normas y rutinas de trabajo que influyen en las relaciones recíprocas de los miembros de una organización y en su colaboración para alcanzar los objetivos". (Jones, Gareth y Jennifer George, 2010: 91).

En el marco de esta investigación se ha optado por un diseño descriptivo experimental y por una metodología cuanti-cualitativa. Para el desarrollo del software se está trabajando con una metodología ágil. Las instancias de evaluación de usabilidad se realizarán con usuarios reales de la institución demandante una vez concluido el proceso de evaluación del prototipo del software, por parte de los especialistas del Centro de Rehabilitación Integral.

4.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 adecuaciones y ajustes en los prototipos. En la ejecución del mismo, se identificaron las funciones ejecutivas y capacidades del grupo destinatario mediante la administración de entrevistas focalizadas al equipo de especialistas -terapeutas del Centro de Rehabilitación Integral.

Los sujetos bajo estudio con los que se está trabajando en Vivitan son niños y adolescentes que asisten a la mencionada institución, cuyas edades cognitivas se encuentran comprendidas entre 10 y 16 años.

Durante la ejecución del proyecto se está adaptando una metodología de desarrollo de software ágil que contemple las características del videojuego, así como también las características singulares de sus destinatarios.

Vivitan pretende generar una tecnología apropiada y de calidad (Berenbach, Paulish, Kazmeier & Rudorfer, 2009, MarkusManies,UolevisNikual Lappeenranta, 2011). Por tal motivo, en una primera instancia, se realizó la elicitación de requisitos mediante: visitas a la institución demandante, entrevistas semiestructuradas a

informantes claves de la Institución adoptante "Centro de Rehabilitación Integral" a fin de conocer la cultura organizacional, su misión, visión, valores y formas de intervención.

En cuanto a su cultura organizacional (Jones, Gareth y George, J. (2010)), el relevamiento realizado a través de las técnicas de recolección de datos permitió conocer la forma en que se desempeñan y se relacionan los miembros con los demás dentro y fuera de la institución. De ahí la importancia de desarrollar valores compartidos para lograr la calidad a la que aspira la dirección.

La Institución demandante tiene como misión la rehabilitación neurológica de los niños y adolescentes que asisten para mejorar su calidad de vida, con terapias adaptadas a cada necesidad específica. El propósito del centro en lograr con eficacia la rehabilitación, mediante las diferentes terapias personalizadas, las cuales están orientadas tanto a niños como adultos, con la participación de profesionales que mediante trabajo interdisciplinario abordan las diferentes dimensiones de la persona. La motivación en el niño y una interacción adecuada entre éste y el especialista contribuyen al éxito del tratamiento.

El equipo de trabajo de la Institución adoptante se reúne y evalúa decisiones y cursos de acción realizados en reuniones interdisciplinarias periódicas, (ateneos) para medir los resultados alcanzados y de esa forma lograr intervenciones exitosas.

En cuanto al software específico, se unificaron criterios para: incorporar sensaciones físicas al mundo virtual dentro de un espacio similar al consultorio, centrándose sólo en la función cognitiva y que dicho desarrollo se adapte a la cultura organizacional de la institución demandante a fin de que sea apropiada y contribuya efectivamente a las intervenciones terapéuticas que se realizan actualmente.

Por otra parte, se realizó un experimento para medir la presencia e inmersión con el sistema de Realidad Virtual implementado, con 16 usuarios jóvenes y adultos voluntarios.

Paralelo a esto, se trabajó en la elaboración de un informe de evidencias, avanzando en la tarea de búsqueda bibliográfica (o narrativa) de herramientas de software aplicadas a parálisis cerebral, en especial las desarrolladas para plataformas móviles. El objetivo perseguido es, junto con el asesoramiento de los terapeutas del Centro de Rehabilitación Integral, efectuar el estudio de algunas de las herramientas que resulten de la búsqueda y llevar a cabo experiencias de aplicación con usuarios con patologías discapacitantes que asisten al Centro de Rehabilitación Integral. Además, se estableció contacto, por un lado, con la empresa ENESO VERBO, especialistas en desarrollo de software para personas con dificultades en la expresión o el lenguaje, de quienes se obtuvo una licencia académica para usar la plataforma eneso.com. y por otro, con grupos de investigación de España entre los que destaca el proyecto FRESSA (https://tecnoaccesible.net/content/projecte-fressa) con quienes se mantiene contacto para el intercambio de experiencias.

El producto tecnológico consiste en un entorno o sistema de realidad virtual, se encuentra en estado avanzado de desarrollo y en etapa de análisis respecto a la mejor forma de lograr la transferencia al Centro de Rehabilitación Integral.

5. 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.

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. Doctora en Educación. Defendió su Tesis denominada"Procesos de Apropiación de TIC en Espacios Socio-educativos no Formales. Propuesta de Formación Integral para Adultos Mayores", el 19 de diciembre de 2018. Doctorado en Educación. Facultad de Educación. Universidad Católica de Cuyo.
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 "*Marco para la gestión de contenidos descargables en Sistemas de Realidad Virtual*".
7. Carrió Fabricio. Trabajo Final de Licenciatura en Ciencias de la Computación en proceso. Denominado "Técnica para el diseño de videojuegos de rol"
8. Kokot Rodrigo: Trabajo Final de Licenciatura, denominado: "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

Muriel V, García-Molina A, Roig-Rovira T. (2014). Estimulación cognitiva en niños con parálisis cerebral. Rev Neurol 2014; 59:443-8

Howcroft J, Klejman S, Fehlings D, Wright V, Zabjek K, Andrysek J, (2012). Active video game play in children with cerebral palsy: potential for physical activity promotion and rehabilitation therapies. Arch Phys Med Rehabil 2012; 93: 1448-56.

González Sánchez J. (2010). *Jugabilidad. Caracterización de la experiencia del jugador en videojuegos* (Tesis doctoral). Universidad de Granada, España.

Jones, Gareth & Jennifer George, (2010) “Administración Contemporánea, Sexta edición, Ed. Mc.Graw-Hill, México.

Dierssen, M. (2014).Videojuego para la estimulación cognitiva de las personas con discapacidad intelectual. https://www.precipita.es/precipitado/videojuego-para-la-estimulacion-cognitiva-de-las-personas-con-discapacidad-intelectual.html

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/

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, Romano Lucas, Blanco Claudio, Aguilera Belén, 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, lromano, cblanco, mbaguilera}@untdf.edu.ar
alejandra.zangara@gmail.com, sergiorc.ush@gmail.com

RESUMEN

El avance de la tecnología ha provocado cambios en el modelo educativo. Como un ejemplo de ello podemos citar los EVEAs (Entornos Virtuales de Enseñanza y de Aprendizaje). Utilizados en sus inicios para resolver los problemas que planteaba la Educación a Distancia (EaD), fueron luego adoptados para complementar la Educación Presencial, dando origen a diferentes modalidades que extienden las posibilidades del aula presencial.

Es importante entonces, que las instituciones educativas, especialmente las universidades, brinden soporte académico a sus docentes y alumnos implementando una plataforma a nivel institucional, ofreciendo así la posibilidad de seleccionar las herramientas que el entorno tiene disponibles, para que cada docente pueda potenciar su propuesta educativa usando estas tecnologías.

En la actualidad existen numerosos EVEAs, de código abierto o comerciales, que cuentan con distintas funcionalidades.

La selección de un EVEA institucional, que responda al modelo educativo de la institución universitaria, 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, para recomendar el que mejor se adecue al modelo pedagógico planteado por la UNTDF [1]

Palabras clave: EVEAs, evaluación de EVEAs, criterios de selección de EVEAs, mediación de la enseñanza

CONTEXTO

El 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). 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). Se desarrolla formalmente desde el 01/03/2017 hasta el 30/04/2019.

1. INTRODUCCIÓN

El avance de la tecnología ha provocado cambios en el modelo educativo, a los que resulta imposible permanecer ajeno. Por ello. muchas instituciones educativas buscan incorporar herramientas que hagan uso de

tecnologías existentes como formas alternativas de comunicación para brindar o mejorar los procesos de enseñanza y de aprendizaje.

Como un ejemplo podemos citar los EVEAs. Utilizados en sus inicios para resolver los problemas que planteaba la EaD, fueron rápidamente adoptados para complementar la Educación Presencial, dando origen a diferentes modalidades que extienden las posibilidades del aula presencial.

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. Entre otras cosas, facilita la comunicación pedagógica entre los participantes en un proceso educativo, sea éste a distancia, presencial o de naturaleza mixta [3].

En la actualidad existe un amplio abanico de EVEAs orientados a la educación superior. Los hay con características de software propietario o libre y en la nube (aunque en este último caso no siempre son considerados como EVEAs propiamente dichas) [4].

Si se consideran los desarrollos propios de distintas universidades, el número de EVEAs supera la centena. Si nos limitamos a los que han logrado cierta preponderancia en el ámbito internacional encontramos al menos 40 de ellos [5][6].

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: [7] 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 permitirá establecer el modelo de enseñanza en el que se basará la capacitación docente y la implementación del entorno seleccionado.

Resulta claro entonces que no debe restringirse la selección solamente a los aspectos técnicos. Al respecto señala Clarenc: "*Es necesario contemplar a los LMS (*Learning Management System) *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*." [8]

Por otra parte, debe tenerse en cuenta que, si bien existen muchos EVEAs de uso gratuito, la inversión que deben realizar las instituciones que los adopten es significativa. Uno de los rubros a los que deben destinarse muchos recursos es la preparación del cuerpo académico para un uso adecuado y eficiente del entorno. Eventualmente podría requerirse el agregado de alguna funcionalidad particular, lo que implica realizar todas las tareas que demanda un

desarrollo de software.

La conclusión es que, más allá de la "supuesta" gratuidad inicial, 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 de cada uno. Mediante esta evaluación podremos determinar su filosofía pedagógica, su flexibilidad didáctica, sus funcionalidades y su usabilidad [3]. Los indicadores más importantes que pueden tenerse en cuenta a la hora analizar un entorno virtual están relacionados con el tipo de actividad educativa a desarrollar (cursos, carreras, comunidades virtuales), la modalidad, las características de los destinatarios y las posibilidades y limitaciones técnicas disponibles (ancho de banda, equipamiento informático y disponibilidad de tiempo, entre otros).

A su vez, 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.

2. LÍNEAS DE INVESTIGACIÓN, DESARROLLO E INNOVACIÓN

Se mencionan aquí las principales líneas de investigación y desarrollo abordadas en el marco del proyecto:

- EVEAs. 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.

3. RESULTADOS OBTENIDOS / ESPERADOS

Dentro del objetivo general del proyecto (seleccionar un EVEA a fin de que sea utilizado como entorno institucional de la UNTDF y contribuir al proceso de implementación) se consideran los siguientes objetivos específicos:

- Analizar las distintas funcionalidades provistas por los EVEA.
- Establecer el conjunto de criterios para la evaluación de los EVEAs.
- Aplicar los criterios a los entornos analizados.
- 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.

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. Se seleccionaron, en una primera etapa, las Plataformas Chamilo, Moodle, Ilias y Sakai.

Previo a su 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, se decidió seleccionar para las experiencias piloto previstas en el proyecto la plataforma Chamilo [9]. Durante el 2º cuatrimestre de 2018, se lo utilizó en materias de 1º, 2º y 3º año de la Licenciatura en Sistemas y en una asignatura de 1º año de las carreras de Contador Público y Gestión Empresarial, que están a cargo de los docentes que participan del proyecto.

Las experiencias han podido confirmar que la elección ha resultado exitosa, en especial por dos factores importantes: su sencillez de uso, justificada en cómo está estructurada cada interfaz con la que cuenta, y su gran atractivo visual por el diseño con el que fue desarrollado.

A partir de las experiencias obtenidas se ha recomendado a la Secretaría Académica ofrecer a todos los docentes de la UNTDF la posibilidad utilizar en sus cursos al EVEA Chamilo (a la fecha sólo se ofrece a nivel institucional la plataforma Moodle).

En caso de obtener una respuesta favorable restará realizar las tareas:

- 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.

4. FORMACIÓN DE RECURSOS HUMANOS

El grupo de trabajo está formado por docentes - investigadores, de los cuales uno (1) es Doctora en Ciencias Informáticas, dos (2) son Especialistas en Docencia Universitaria, dos (2) son Licenciados en Informática, uno es Licenciado en Administración y Sistemas, uno (1) es Analista Programador Universitario y uno (1) es alumno avanzado de la carrera Licenciatura en Sistemas.

Uno de ellos se encuentra cursando la maestría en Tecnología Informática Aplicada en Educación en la Facultad de Informática de la UNLP. Otro está cursando la maestría en Ingeniería de Software en la Facultad de Informática de la UNLP.

El alumno Rojas desarrolló 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 4 de marzo de 2018). Estatuto de la UNTDF [Online]. Disponible en www.untdf.edu.ar/reglamentacion.

[2] Patricia Avila M., Martha Diana Bosco H.. (2001. Accedido el 28 de febrero de 2019). 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. Accedido el 28 de febrero de 2019) Hacia un modelo de evaluación de entornos virtuales de enseñanza y aprendizaje. La importancia de la usabilidad. Revista TEyET [Online] Nº 4 pp. 10 - 14. Disponible en: tp://teyet-revista.info.unlp.edu.ar/numero-4-octubre-de-2009/

[4]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 4 de marzo de 2019

[5] Baumgartner P, Häfele H., et al. E-Learning Praxishandbuch: Auswahl von ernplattformen. Marktübersicht - Funktionen Fachbegriffe.Innsbruck-Wien, StudienVerlag. 2002.

[6] Adell J. (2004. Accedido el 4 de marzo 2019) 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

[7] Adell J. (2004. Accedido el 4 marzo 2019) 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

[8] 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-.

[9] Plataforma Chamilo LMS: ¡Las novedades de 1.11! - Slideshare Disponible en https://es.slideshare.net/chamiluda/plataforma-chamilo-lms-las-novedades-de-111 Accedido el 4 de Marzo 2019

Implementación de una Plataforma de Capacitación a Distancia Abierta y Accesible con Recursos Educativos destinados a Personas con Discapacidad Visual

Guillermo Javier Lafuente[1], Carlos Ballesteros[2], José Luis Filippi[3], Gustavo Hernán Lafuente[4]

GIAU[5] – Facultad de Ingeniería – UNLPam.
Calle 110 esq. 9 nº 390
{lafuente[1], balleste[2], filippij[3],gustavo[4]}@ing.unlpam.edu.ar
[5]Grupo 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 tiene como finalidad 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. Las capacitaciones serán dictadas por la Fundación BienEstar[1], especializada en el trabajo con personas con discapacidad.

Para contribuir a la resolución del proyecto, en la actualidad se ha implementado un EAD (Entorno de Educación a Distancia) que contempla las características de Accesibilidad explicitadas en el (Artíc. 9° de la Convención Internacional de los Derechos de las Personas con Discapacidad Ley (26.378). En forma paralela, se encuentra en proceso de desarrollo un conjunto de Recursos Educativos Abiertos y Accesibles (REAAs), con el objetivo de ser instrumentos de capacitación en el contexto del EAD implementado, en el marco de alguna disciplina 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 dentro 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 involucra el diseño e implementación de una plataforma de capacitación a distancia utilizable y abierta para personas con discapacidad visual. El concepto de diseño universal

[1] Fundación BienEstar, Realicó, La Pampa http://www.fundacionparaelbienestar.org/

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 termino accesibilidad se ha extendido al ámbito de la informática. Junto al crecimiento de Internet, también crecieron estrategias de accesibilidad accesibilidad para la Web. Diversos autores [4] [5] han desarrollado este tópico, describiendo tecnologías que asisten a 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 nuevos desafíos que conlleva a involucrar nuevas técnicas para la 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 un 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: Por consiguiente, una buena experiencia de usuario para el caso de ciegos, no está para nada asegurada. De acuerdo a esta premisa, se necesitan aplicaciones que contemplen a usuarios con baja visión o no videntes, logrando una mejor experiencia de usuario posible, con un Desarrollo Centrado en este tipo de Usuario (DCU).

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. Según Piñeros [11], la desventaja radica, 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 capacitar 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 está siendo abordado mediante la metodología Design Thinking [13], combinadas junto a la metodología Lean Startup [14].

Desde el punto de vista de la accesibilidad en los productos desarrollados, se están implementando 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 elaboraron 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 produce actualmente es evaluado para alcanzar el nivel máximo de accesibilidad. Para ello, se utiliza un conjunto de herramientas automatizadas [15], [16], [17] que asisten a los procesos de desarrollo, para obtener un mayor grado de accesibilidad sobre los productos construidos. Estas herramientas examinan la accesibilidad bajo las directrices de accesibilidad web WCAG, e informan sobre las barreras conocidas y probables o potenciales (en forma de un informe independiente y / o anotaciones del producto evaluado).

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 [15], [16], [17] son agrupados y tabulados para preparar una lista única de problemas potenciales de accesibilidad para cada sección de los artefactos de software. Se involucra en este proceso cada unidad de información contenida en los REAAs elaborados.

b) *Estrategia de evaluación de accesibilidad centrada en el estudiante/usuario*

En el futuro, 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 colaboraran 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 [18] y Rubin y Chisnell [19]. 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 [20].

Para relevar estas pruebas de laboratorio se utilizarán herramientas para captura de escenas como OpenVULab o Screencast-O-Matic[2], 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.

[2] OpenVULab (http://openvulab.org) Screencast-O-Matic (https://screencast-o-matic.com/)

Líneas de Investigación, Desarrollo e Innovación

A continuación, se describen sintéticamente las líneas de investigación en las cuales se avanza 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 y accesible, y un REAA a través de un diseño centrado en usuarios con discapacidad visual, que esté diseñada y probada por este tipo de usuarios, haciéndolos partícipes del proceso de desarrollo conforme a las metodologías de trabajo implementadas.

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/. [Accessed: 14-Mar-2018].

[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/. [Accessed: 06-Oct-2017].

[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] B. Tim, "Design Thinking By Tim Brown," *Harv. Bus. Rev.*, 2008.

[14] E. Ries, *The Lean Startup: How Today's Entrepeneurs Use Continuous Innovation to Create Rdically Sucesful Businesses*. 2011.

[15] "WAVE Web Accessibility Tool." [Online]. Available: http://wave.webaim.org/. [Accessed: 16-Mar-2018].

[16] "IDI Web Accessibility Checker : Web Accessibility Checker." [Online]. Available: https://achecker.ca/checker/index.php. [Accessed: 16-Mar-2018].

[17] "Home: AInspector Sidebar." [Online]. Available: http://ainspector.github.io/. [Accessed: 16-Mar-2018].

[18] J. Nielsen, *Usability Engineering*, vol. 44, no. 3. 1993.

[19] D. Rubin, J., & Chisnell, *Handbook of usability testing. How to plan, design, and conduct effective tests (2nd ed.)*. 2008.

[20] M. T. Boren and J. Ramey, "Thinking aloud: Reconciling theory and practice," *IEEE Trans. Prof. Commun.*, vol. 43, no. 3, pp. 261–278, 2000.

Integración de recursos del Paradigma Analítico y de la Inteligencia de Negocios como estrategia para el fortalecimiento del proceso de toma de decisiones en un contexto educativo.

Ing. Corso Cynthia, Ing. Constable Leticia, Ing. Colaccioppo Nicolás, Ing. Agustina Cragnolini, Mallo Britos Anabel.

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/leticiaconstable@gmail.con/nicolas_colacioppo@hotmail.com/agustina cragnolini@gmail.com/anabelmbritos@gmail.com

RESUMEN

El desarrollo de las tecnologías de información y comunicación han facilitado el uso de herramientas tecnológicas en numerosos ámbitos. Esto ha propiciado la generación de información en volúmenes significativos, que dificulta la posibilidad de extraer decisiones basadas en situaciones detectadas como no deseables. Es por eso que las organizaciones necesitan la integración de herramientas tecnológicas adecuadas para poder adaptarse a las exigencias del medio y detectar de manera más ágil las posibilidades de mejora. Por lo tanto este proyecto pretende el desarrollo de una metodología para el diseño de un sistema de soporte decisión que será aplicado en un contexto educativo.

.

Palabras claves: *inteligencia de negocios, paradigma analítico, sistema de soporte de decisión, educación.*

CONTEXTO

Este trabajo hace referencia al proyecto "Integración de recursos del Paradigma Analítico y de la Inteligencia de Negocios como estrategia para el fortalecimiento en el proceso de toma de decisiones" PID-SIUTNCO0005101, que ha sido homologado por la Secretaría de Ciencia, Tecnología y Posgrado de la Universidad Tecnológica Nacional. El contexto de desarrollo de la presente investigación es el Centro de Investigación, Desarrollo y Transferencia de Sistemas de Información *(GIDTSI)* radicado en la U.T.N Facultad Regional Córdoba.

1. INTRODUCCIÓN

Actualmente las instituciones educativas enfrentan un contexto con mayores demandas y exigencias, por lo que resulta importante tener la posibilidad de tomar decisiones inteligentes, colaborativas y basadas en hechos para conseguir resultados óptimos.

Existen diversos estudios que han considerado el uso de analíticas como un recurso válido para fortalecer el proceso de toma de decisiones de manera exitosa en diferentes ámbitos [1][2][3][4]. La capacidad analítica consiste en la utilización de herramientas y tecnologías que convierten los datos en utilizables, intuitivos y predictivos. Este concepto puede llegar a potenciarse por medio del

uso de herramientas pertenecientes al área de la inteligencia de negocios, como una alternativa para el fortalecimiento en el proceso de toma de decisiones.

La Inteligencia de Negocios (BI-Business Intelligence) comprende un conjunto de estrategias y componentes que permiten transformar los datos operacionales en información consolidada y ésta en conocimiento útil para la toma de decisiones acertadas; es decir, facilita el monitoreo del cumplimento de los objetivos organizacionales y admite el análisis de la información histórica, contribuyendo la creación de estrategias que generan ventajas competitivas en el mercado [5].

Los campos de aplicación de BI más comunes son: educación, el sector empresarial e industrial, la banca, salud, turismo, moda y otros, en donde se pueda analizar las tendencias e identificar la demanda de los consumidores [6][7][8].

Existen diversas herramientas de BI que facilitan el análisis descriptivo y predictivo de los datos. Este tipo de herramientas se clasifican en dos: los sistemas de soporte de decisiones y los sistemas de información para ejecutivos [9][10][11][12].

Los sistemas de soporte de decisiones (SSD) se caracterizan por modelar la información usada por los altos mandos como apoyo para la toma de decisiones [13]. Estos facilitan el análisis de datos debido a que disponen de información histórica accesible rápidamente y presentada en informes dinámicos, flexibles e interactivos.
Mientras que los sistemas de información ejecutiva (EIS) tienen la función de analizar el estado actual de los KPI's de cada área de la organización, manejando información estratégica para la empresa.

Este trabajo se enfoca en el estudio relacionado con el diseño e implementación de herramientas de inteligencia de negocio, cuyos resultados serán transferidos a un contexto educativo de nivel superior. La fundamentación de transferencia de los resultados a este ámbito, radica en la importancia de poder encontrar nuevos recursos alternativos que faciliten el mejoramiento en materia de calidad educativa basado en la toma de decisiones inteligentes.

La unidad bajo estudio es la cátedra de Paradigmas de Programación perteneciente al segundo nivel de la carrera de Ingeniería en Sistemas de Información de la UTN Facultad Regional Córdoba. Desde la cátedra se promueve la incorporación de herramientas tecnológicas que permitan contribuir a una educación de calidad a través del asesoramiento y acompañamiento de las actividades académicas desarrolladas. Una de las herramientas que se encuentran a disposición de los docentes son los entornos virtuales de enseñanza y aprendizaje (EVEA), cuyas posibilidades de utilización son muchas y muy variadas [14]. La UTN Facultad Regional Córdoba ofrece la posibilidad de acceder a la plataforma virtual Moodle, que representa un espacio de trabajo para los profesores y alumnos facilitando la comunicación, interacción y acceso a recursos de interés.

La cátedra bajo estudio ha manifestado la necesidad de disponer información consolidada e integrada relacionada con el progreso y resultado de las actividades académicas desarrolladas de manera presencial y on-line que surgen de la interacción con la plataforma virtual

Moodle. Por lo expuesto anteriormente se infiere que el tomar decisiones en este contexto, en el que el volumen de datos es significativo se torna en una actividad sumamente compleja para los profesores. Dando origen así a la necesidad de desarrollar sistemas de apoyo para la toma de decisiones que se proponen como objetivo del presente proyecto.

2. LINEAS DE INVESTIGACIÓN y DESARROLLO

De este proyecto se desprenden al menos cuatro líneas de investigación y desarrollo. A continuación se presenta una descripción sintética de las mismas:

Inteligencia de Negocios: se pretende el estudio y análisis de herramientas de BI haciendo foco en los sistemas de soporte de decisión, su arquitectura y la revisión metodologías para su diseño e implementación.

Ingeniería de Requerimientos: se investigarán diversas técnicas de requerimientos con el propósito de formalizar y registrar las demandas del dominio bajo estudio.

Datawarehouse: se estudiarán distintos enfoques para el diseño del almacén de datos, como así también la evaluación de estrategias para la integración de los datos que provienen de diferentes fuentes.

Paradigma Analítico: en este proyecto se considerará el estudio de las analíticas de aprendizaje cuya finalidad es el seguimiento del uso de los entornos virtuales. El estudio de este aspecto permitirá complementar los resultados de esta investigación en el contexto bajo estudio.

Entornos virtuales de enseñanza y aprendizaje: se investigarán y analizarán herramientas de seguimiento para la recolección de los datos que surgen de las interacciones de los alumnos a las actividades y con los profesores.

3. RESULTADOS OBTENIDOS/ESPERADOS

Este proyecto se ha planificado en tres años de ejecución (2019-2021) para el desarrollo de los siguientes objetivos:

Objetivo General

Desarrollar una metodología para el diseño e implementación de un sistema de información que agilice el proceso de toma de decisiones, considerando la analítica digital e indicadores de interés para el dominio bajo estudio.

Objetivos Particulares

- Relevar la situación actual del dominio bajo estudio e identificar los aspectos de los que se pretende realizar un seguimiento para su control y optimización.
- Definir objetivos y métricas analíticas alineadas con los requerimientos detectados.
- Elaborar un esquema que permita implementar un almacén con el fin de consolidar los datos de interés que provienen de diferentes fuentes.
- Diseñar y desarrollar una interfaz que permita la integración de los datos almacenados para la detección de desviaciones de los objetivos propuestos en el dominio de análisis.

Resultados esperados

Como este proyecto está al inicio de su primer año de ejecución los resultados que se han logrado hasta el momento

están vinculados con el estado del arte de las herramientas tecnológicas que pertenecen al campo de la inteligencia de negocios.

A continuación se detallan las actividades planificadas para el desarrollo del proyecto:

- Relevamiento a los profesores de la cátedra bajo estudio con el fin de establecer la misión, visión y estrategia en este contexto.
- Diseño del mapa estratégico atendiendo las necesidades de la cátedra alineado con la misión de la misma.
- Identificación de los indicadores para cada objetivo identificado.
- Relevamiento de herramientas tecnológicas para la automatización del sistema de soporte de decisión.
- Implementación del sistema de soporte de decisión.
- Validación de funcionalidad con los usuarios finales.
- Capacitación dirigida a los usuarios.
- Realización de revisiones estratégicas periódicas, con el fin de obtener un feedback para aprender sobre la estrategia y mejorarla.

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.

La mayoría de los docentes investigadores que integran el PID, comparten su actividad académica en la cátedra Paradigmas de Programación lo que constituye un aporte significativo para el avance y ejecución del proyecto.
Además se prevé la participación de 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 dos becarios, un alumno y otro graduado con el objetivo de complementar su formación académica con un acercamiento al ámbito de la investigación científica.

5. REFERENCIAS

[1] Siemens, G. Learning analytics: envisioning a research discipline and a domain of practice. En Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, ACM, pp. 4-8, 2012.
[2] Clow, Doug. The learning analytics cycle: closing the loop effectively. National Conference on Learning Analytics & Knowledge, Vancouver, pp. 134-137, 2012.
[3] Baker, R.; Siemens, G. Educational data mining and learning analytics, Sawyer, K. (ed.), Cambridge Handbook of the Learning Sciences, 2013.
[4] Marín Díaz, V.; Sampedro Requena, B. Web 2.0 for the invigoration and participation of families and communities, Journal of New Approaches in Educational Research, vol.5, pp. 38-43, 2016.
[5] Marinheiro, A., & Bernardino, J. Analysis of open source Business Intelligence suites. Information Systems and Technologies (CISTI), 2013 8th Iberian Conference on, pp. 1-7, 2013.

[6] Parr, O. Data Mining Cookbook Modeling Data for Marketing, Risk, and Customer Relationship Management, Editorial Wiley, 2000.
[7] Silberschatz A., Korth H., Sudarshan S. Fundamentos de Base de Datos, McGraw-Hill, Madrid, España, 2006.
[8] Abebe, A. J. and Price, R. K. Decision support system for flood warning in urban areas. Proceedings of the XXIX IAHR Congress, Beijing, China, Sept 16-21, 2001.
[9] Ghosh, R. An Integrated Approach to Deploy Data Warehouse in Business Intelligence Environment, Third International Conference on Computer, Communication, Control and Information Technology (C3IT 2015), ISBN: 978-1-4799-4445-3, 2015.
[10] M. S. Gounder, V. V. Iyer and A. A. Mazyad, A survey on business intelligence tools for university dashboard development, 2016 3rd MEC International Conference on Big Data and Smart City (ICBDSC), Muscat, pp. 1-7, 2016.
[11] Marinheiro, Antonio & Bernardino, Jorge. Experimental Evaluation of Open Source Business Intelligence Suites using OpenBRR. Latin America Transactions, IEEE (Revista IEEE America Latina). 13. 810-817, 2015.
[12] M. Vassell, O. Apperson, P. Calyam, J. Gillis and S. Ahmad, "Intelligent Dashboard for augmented reality based incident command response co-ordination," *2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, NV, pp. 976-979, 2016.
[13] Castro, C., Uribe, D., & Castro, J. Marco de Referencia para el Desarrollo de un Sistema de Apoyo para la Toma de Decisiones para la Gestión de Inventarios, 10(1), pp. 32–44, 2014.
[14] Nóbile Cecilia Inés., Luna Álvaro Enrique. Los entornos virtuales de Enseñanza y Aprendizaje en la Universidad Nacional de La Plata. Una aproximación a los usos y opiniones de los estudiantes. INNOEDUCA, Editorial: Universidad de Málaga,Vol. 1, pp. 3-9, 2015.
[15] Dawson, S., Heathcote, L. & Poole, G. Harnessing ICT potential: The adoption and analysis of ICT systems for enhancing the student learning experience. International Journal of Educational Management, pp. 116-128, 2010.
[16] Andreu J., Solera, A. and Paredes, J. Decision support system for integrated water resources planning and management. Northern Environmental Research Network, University of Oulu, Finland, 2002.
[17] Quintero, J. H. Diseño de un modelo gerencial basado en el cuadro de mando integral para el Instituto Universitario Tecnológico de Ejido, Actualidad Contable Face, pp. 89-103, 2008.
[18] Rodríguez-Cruz, Y.; Pinto, M. Modelo de uso de información para la toma de decisiones estratégicas en organizaciones de información. Transinformação, vol. 30, pp. 51-64, 2018.
[19] Barra, A. M., Alineamiento Estratégico Sectorial: Caso de Estudio Aplicado a una Universidad Chilena, Formación Universitaria, pp. 3-12, 2015.
[20] Verbert, K.; Govaerts, S.; Duval, E.; Santos, J. L.; Van Assche, F.; Parra, G.; Klerkx, J. Learning dashboards: an overview and future research opportunities, Personal and Ubiquitous Computing, pp. 1499-1514, 2014.

La ubicuidad presente en una Comunidad virtual para docentes de Matemática de Zonas Alejadas

Liliana Rios(1), Rosa María Pósito, Patricia Cuadros(2), Vanesa Gallardo, Sonia Capdevila, Inés Calvo, Victorina Márquez, Lucía Gil
Facultad de Ciencias Exactas Físicas y Naturales
CUIM – Av. Ignacio de la Roza 590 (O), Rivadavia – J5402DCS San Juan, 0264 4265101
(1) Facultad de Filosofía, Humanidades y Artes
(2) Facultad de Ingeniería
{lrselene, rosaposito, pcuadrosan, ep2012gallardovanesa, secapdevila}@gmail.com
{icalvo, vmarquez,lgil}@unsj-cuim.edu.ar

RESUMEN

Los estudiantes de la educación secundaria y de los primeros años de la universidad, llamados "Los Millennials" acceden al mundo mediante las TIC desde su alfabetización. Tienen acceso instantáneo a la información y esto genera un nuevo paradigma que influye directamente en el quehacer docente. Surge así la necesidad de los docentes de acceder a propuestas de formación y/o actualización en condiciones equitativas y sostenidas en el tiempo, para lograr una adecuada mediación pedagógica y comunicacional de los contenidos curriculares, utilizando recursos tecnológicos.

El proyecto que se presenta aborda la problemática relacionada con la capacitación actualizada de los docentes de Matemática del nivel secundario de Zonas alejadas, con el fin de promover en los alumnos de dichas zonas la continuidad en los estudios universitarios y evitar la deserción una vez insertos en el sistema universitario, priorizando la igualdad de oportunidades.

Este artículo comparte los avances en la investigación de la implementación de la comunidad virtual de práctica ,a la que hemos denominado CoVMat destinada a docentes de Matemática de zonas alejadas al Gran San Juan; éstatiene la particularidad de ser un espacio virtual que no tiene limitaciones de permanencia en el tiempo, pues una vez terminada esta propuesta de actividad el espacio sigue activo con la posibilidad de continuación,

Se espera que este espacio virtual constituya un punto de encuentro de problemáticas e inquietudes de los docentes tanto de nivel medio como universitario y pueda contribuir a la problemática de deserción de los alumnos de zonas alejadas.

Palabras clave: *Comunidad Virtual, Ubicuidad, Docentes de zonas alejadas, Matemática*

CONTEXTO

El artículo que se presenta pertenece al proyecto de investigación "La ubicuidad de una comunidad virtual como estrategia para acceder al perfeccionamiento de profesores de matemática de educación secundaria de zonas alejadas", Código: 21/E1057 . Este proyecto está inserto en la línea de Tecnología Informática Aplicada en Educación, del Programa de Investigación del Gabinete de Tecnología e Innovación Educativa GATIE de la Facultad de Ciencias Exactas Físicas y Naturales de la UNSJ. Fue presentado por docentes de la universidad que persiguen acercar herramientas para mejorar el acceso de docentes de Matemática de la educación secundaria que ejercen su actividad en escuelas de zonas alejadas al gran San Juan a ofertas de

actualización tanto tecnológica como didáctico pedagógicas.

1. INTRODUCCION

La Tecnología Educativa (TE) ha sido una de las disciplinas de la Ciencias de la Educación que más ha evolucionado en los últimos años. Su concepto se ha ido modificando de acuerdo a dos grandes visiones: la aplicación de los medios y el diseño de la instrucción (Cabero, 2001 y Cabero, Barroso y Llorente, 2015). Cabero (2001) sostiene que el campo de acción de la TE es el diseño de situaciones de aprendizaje mediadas con TIC, donde se pone énfasis en los procesos seguidos y no en los productos finales. Estos procesos deben adaptarse a los contextos concretos de aplicación. La formación educativa con tecnología tiene que estar acompañada de propósitos y valores social e individualmente relevantes, que responda a las demandas educativas del sistema, en particular en este proyecto se aborda la problemática de los docentes de Matemática de zonas alejadas y las dificultades que tienen para acceder a una capacitación sostenida en el tiempo.

La actualización permanente del profesorado es primordial para un uso oportuno de las TIC en los procesos de enseñanza y aprendizaje. Pero los docentes de zonas alejadas a los centros urbanos tienen inconvenientes para acceder a propuestas de formación y/o actualización. Desde esta perspectiva la visión de Nicholas Burbules trae luz a esta problemática al afirmar que:

> El futuro de la formación docente tendrá que abordar el aprendizaje ubicuo: la posibilidad de acceder a la información en cualquier lugar o cualquier momento, la interacción con pares y expertos eruditos y oportunidades estructuradas de aprendizaje desde una variedad de fuentes. La brecha entre el aprendizaje formal e informal desaparecerá.

En el contexto social actual, atravesado por la el papel del profesor sigue siendo crucial, pero configurado como mentor, facilitador, modelo a seguir, o tutor. Para llevar a cabo estas funciones de buena manera en un contexto de aprendizaje ubicuo, los docentes tienen que estar mejor informados y más comprometidos con lo que los alumnos ya saben. Ello requiere potenciar las competencias digitales del profesorado. Desde este proyecto se pone énfasis en que las competencias y destrezas se refieren a como conocer y comprender (conocimiento teórico de un campo académico, la capacidad de conocer y comprender), saber cómo actuar (la aplicación práctica y operativa del conocimiento a ciertas situaciones) y saber cómo ser (los valores como parte integrante de la forma de percibir a los otros y vivir en un contexto social).

Diversas organizaciones, entre ellas el Parlamento Europeo han realizado recomendaciones y consensuado definir sobre ocho competencias claves que cualquier joven debe haber desarrollado al finalizar la enseñanza obligatoria para poder incorporarse a la vida adulta de manera satisfactoria y ser capaz de desarrollar un aprendizaje permanente a lo largo de la vida.

(Recomendación 2006/962/CE del Parlamento Europeo y del Consejo, de 18 de diciembre de 2006, sobre las competencias clave para el aprendizaje permanente, Diario Oficial L 394 de 30.12.2006).

De estas ocho competencias se aborda especialmente en este trabajo de investigación, la competencia del tratamiento de la información y competencia digital -entendida como la capacidad del alumno para buscar, obtener, procesar y comunicar información y transformarla en conocimiento-; y la competencia Matemática que supone la habilidad para utilizar y relacionar números, sus operaciones básicas y el razonamiento Matemático para interpretar la información, ampliar el conocimiento y resolver problemas

tanto de la vida cotidiana como del mundo laboral.

Jordi Adell (2008) define la competencia digital como una de las 8 básicas que deben formar parte de la educación básica y secundaria y afirma que es necesaria para aprender a lo largo de toda la vida, para que alumno alcance el "aprender a aprender".

Para la Unesco "Un docente que no maneje las tecnologías de información y comunicación está en clara desventaja con relación a los alumnos. La tecnología avanza en la vida cotidiana más rápido que en las escuelas, inclusive en zonas alejadas y pobres con servicios básicos deficitarios."

Respecto a la competencia, la Matemática ha recibido el impacto de la introducción de las computadoras y otros tipos de tecnologías, como las calculadoras gráficas, que requieren cambiar algunas cuestiones relacionadas con la enseñanza de los contenidos, dado que su gran capacidad y rapidez en el cálculo, y la facilidad que brindan para lograr representaciones gráficas, permiten incursionar aún más en otros procesos mentales como la reflexión y la demostración de hipótesis. Son herramientas que sirven para complementar y agilizar tareas cotidianas, desplegando un abanico de posibilidades que conducen al desarrollo de nuevos modelos pedagógicos, como lo es, el aprendizaje colaborativo.

En este marco, la Educación asume un rol fundamental en la formación de competencias para esta nueva sociedad en donde la información fluye por diferentes canales, en forma incesante y es el centro de las actividades productivas. Además se ve interpelada en tanto el conocimiento está disponible en diferentes ámbitos y soportes lo cual lleva a referir al "aprendizaje ubicuo" (Burbules, 2009).

Cabero (2011) señala que los nuevos escenarios son cada vez más ricos en multimedias interactivas y deslocalizadas del espacio y del tiempo suponen:

"Cambio de paradigma: pasar de la transmisión de la información a la construcción del conocimiento; cambia el rol del docente: es guía orientador y diseñador; trabajar con nuevas herramientas de comunicación; creación de entornos de formación activos; trabajar en una cultura de la colaboración donde el intercambio de la información y la construcción conjunta de conocimientos se convierten en la pieza clave de desarrollo; creación de entornos de formación activos creación de entornos personales de formación…"

Este contexto conlleva a la necesidad de propuestas de formación amplias que aborden la planificación, diseño y evaluación y fomenten la coproducción de materiales entre profesores y estudiantes.

En este sentido Mishra&Koehler (2006 y 2008) han propuesto un modelo de análisis del funcionamiento de la TIC en los procesos de enseñanza aprendizaje, denominado TPACK TechnologicalPedagogical Content Knowledge y con él pretenden explicar el conocimiento que debe tener el profesor para la integración de las TIC.

Este modelo parte de la asunción de que todo profesor necesita capacitarse en tres grandes componentes a la hora de incorporar las TIC en los procesos de Enseñanza Aprendizaje: disciplinar/epistemológico, pedagógico y tecnológico. Asumiendo que tales componentes deben verse en interacción, el diseño de una clase con TIC involucra la toma de decisiones epistemológicas /curriculares, pedagógicas y tecnológicas, fundamentadas en criterios, en expectativas de logro, en conocimientos y en experiencias previas. (Rios, 2015).

2. LINEAS DE INVESTIGACION Y DESARROLLO

Este trabajo se encuentra en el marco del Proyecto de Investigación y Desarrollo "La ubicuidad de una comunidad virtual como estrategia para acceder al perfeccionamiento de profesores de matemática de educación secundaria de zonas alejadas", dentro del área temática de Informática Aplicada en Educación. El plan de actividades corresponde al proceso de investigación cualitativa de tipo proyectiva, que implica explorar, describir y proponer acciones, con objetivos de corto, mediano y largo plazo, y una duración prevista de dos años 2018-2019.

Las tareas previstas para el Primer Año fueron:

- Relevamiento de los datos relacionados a las escuelas interesadas en participar de la experiencia.
- Evaluación del grado de acceso a Internet y uso de TIC de los profesores de matemática de escuelas secundarias de zonas alejadas
- Evaluación de las competencias digitales de los docentes de matemática de escuelas secundarias de zonas alejadas
- Conformación de una comunidad virtual de docentes de matemática que pertenecen a escuelas secundarias de zonas alejadas, implementada en el campus de la UNSJ
- Identificación y selección de temáticas de interés, del área matemática, para los participantes de la experiencia, cuyo abordaje se realiza en el seno de la comunidad virtual.

Actualmente se encuentra en desarrollo el Segundo año en el que se preveen las siguientes acciones:

- Monitoreo del funcionamiento de la comunidad virtual y evaluación de fortalezas y debilidades.
- Talleres
- Analisis del rendimiento academico de los alumnos de curso de ingresos a la Facultad de C. E. y N que provienen de zonas alejadas
- Identificación de los errores más comunes de los alumnos ingresantes, a fin de ser trabajados con los docentes en la comunidad
- Validación de los resultados obtenidos en la aplicación de la experiencia
- Elaboración de informes y publicaciones.

Las tareas de relevamiento se realizaran en las instituciones a las que pertenecen los docentes de matemática de zonas alejadas al Gran San Juan. Los talleres se realizarán en escuelas seleccionadas por departamento y en la comunidad virtual.

3. RESULTADOS OBTENIDOS

Dentro de las posibilidades de transferencia se encuentran:

- Transferencia de conocimientos y experiencias, a través de talleres y jornadas de difusión a docentes de matemática de escuelas secundarias de zonas alejadas.
- Publicaciones en congresos, revistas nacionales e internacionales.
- Participación en eventos de nivel científico-académico.

4. FORMACION DE RECURSOS HUMANOS

Este proceso de investigación propicia un espacio para la construcción del conocimiento y el desarrollo de experiencias, favorecido por la ubicuidad que plantea el uso de una comunidad virtual para la producción de propuestas didácticas aplicables en el proceso de enseñanza-aprendizaje de la Matemática, que favorece tanto al equipo de investigación; a los alumnos de los profesorados de la UNSJ que muestren interés en la temática; a los docentes de matemática de la educación secundaria de zonas alejadas, agentes multiplicadores de esta metodología de enseñanza con el uso de las TIC y a la comunidad educativa ofreciéndole el

perfeccionamiento y proporcionando material didáctico que les permita utilizarlo en forma independiente.

5. BIBLIOGRAFIA

• Adell, Jordi. (2008) Algunas ideas sobre cómo desarrollar la competencia digital en Primaria y ESO. Centro de Educación y Nuevas Tecnologías Universitat Jaume I. Disponible en:

http://www3.uji.es/~jpuig/Jordi_Adell%20.pdf

• Burbules N. (2009) Entrevista Parte I y II: "Los problemas no se solucionan con prohibir las TIC, simulando que no existen. Las nuevas tecnologías son herramientas demasiado valiosas como para dejarlas fuera del aula", disponible en: http://portal.educ.ar/noticias/entrevistas/nicholas-burbules-los-problema-1.php

• Burbules N. y Callister T. (2006). Educación: Riesgos y Promesas de Las Nuevas Tecnologías de la información. Argentina: Granica.

• Burbules N. (2012). El aprendizaje ubicuo y el futuro de la enseñanza. Revista Rencontres on Education Vol. 13, 2012, 3 – 14.

• Cabero, J. (2016). ¿Qué debemos aprender de las pasadas investigaciones en Tecnología Educativa?. RIITE. Revista Interuniversitaria de Investigación en Tecnología Educativa, 0, pp. 23-33. Doi:

http://dx.doi.org/10.6018/riite/2016/256741.

• Cabero, Julio. (2001). Tecnología educativa. Diseño y utilización de medios en la enseñanza. Barcelona: Paidós.

• Las competencias Básicas en Educación (PDF). Programa Nacional de Formación Permanente Nuestra Escuela. En Recursos del Ministerio de Educación de San Juan. Disponible en: http://www.sanjuan.edu.ar/mesj/LinkClick.aspx?fileticket=Erbkud6SpCM%3d&tabid=570.

• Poyecto Tunning. Disponible http: //www.unideusto.org/tuning/

• Rios, Liliana G. (2012). La comunidad virtual de práctica: un espacio de colaboración y reflexión para docentes de matemática. UNLP. Disponible: http://hdl.handle.net/10915/23572.

• Tobón, S. (2006). Aspectos básicos de la formación basada en competencias. Talca: Proyecto Mesesup. Recuperado de http://www.uv.mx/facpsi/proyectoaula/documents/Lectura5.pdf

Personalización y adaptación en un ambiente virtual de aprendizaje basada en estilos, conocimiento previo y errores frecuentes

González Marcela [(2),] Benchoff Delia Esther[(1) (2)], Constanza Huapaya[(1) (2)], Guillermo Lazurri[(1)], Leonel Guccione[(1)] y Franscisco Lizarralde[(1) (2)]
[(1)]Grupo de Investigación en Inteligencia Artificial Aplicada a Ingeniería / Departamento de Informática/ Facultad de Ingeniería
Juan B. Justo 4302, 0223 4816600 int 259
[(2)] Instituto de Psicología Básica, Aplicada y Tecnología. IPSIBAT (UNMdP – CONICET)
Facultad de Psicología
Complejo Universitario - Funes 3250 - Cuerpo V, T.e. 0223) 475- 2526
Universidad Nacional de Mar del Plata
mpgonza@mdp.edu.ar, benchoff@mdp.edu.ar, {constanza.huapaya, guillesky, leonel.guccione, francisco.lizarralde}@gmail.com

RESUMEN

Las características de los estudiantes al aprender pueden verse reflejadas en sus estilos de aprendizaje y en sus conocimientos previos. Los avances en el uso de las tecnologías de la información aplicadas a la educación, toman en cuenta esta heterogeneidad, facilitando la personalización del aprendizaje mediante la adaptación de ambientes virtuales de aprendizaje (AVAs). Así, pueden diseñarse estrategias de enseñanza con mayor adecuación a la diversidad, y dirigidas a mejorar el rendimiento. Este proyecto, en línea con trabajos anteriores, se propone ampliar y profundizar el proceso de personalización e incrementar el nivel de adaptación en un AVA, dirigido a estudiantes universitarios, partiendo de la focalización en los estilos cognitivos, conocimiento previo y errores frecuentes establecidos en un dominio. En función de la evaluación diagnóstica de los estudiantes, se emplearán los recursos adaptativos que provee Moodle, para diseñar itinerarios de aprendizaje, materiales de estudio y actividades personalizadas, que respondan a las características diagnosticadas. Se explorará la selección de itinerarios y actividades que realicen los alumnos, en relación a sus estilos. La actividad en el AVA será evaluada formativamente y comparada con el rendimiento en exámenes parciales.

Palabras clave: personalización del aprendizaje, diseños adaptativos, estilos de aprendizaje, Ambiente Virtual de Aprendizaje.

CONTEXTO

Nuestra línea de I/D está inserta en el proyecto *Personalización en un ambiente virtual de aprendizaje basada en estilos, conocimiento previo y errores frecuentes,* orientada a estudiantes de ingeniería y psicología.
Esta línea continua las indagaciones ya iniciadas en un proyecto anterior denominado: *Adaptación en un ambiente virtual de aprendizaje: pruebas y materiales personalizados* (ING464/16), y además, la temática está incluida en un proyecto mayor en desarrollo titulado: *Indagaciones para el diseño y desarrollo de propuestas de formación sostenedoras de la motivación en estudios dentro del continuo presencial - distancia: entre derechos, emociones, memes, estilos, estrategias cognitivas, y personalizaciones de y para los estudiantes del nivel superio*r (PSI276/18).

1. INTRODUCCIÓN

Los estudiantes, como todos los individuos, aprenden de manera diferente, poseen conocimientos previos y experiencias educativas diversas, y manifiestan distintos intereses, necesidades y expectativas en su proceso de aprendizaje. Desde la docencia, conocer y comprender esta diversidad, a partir

de la identificación del perfil cognitivo del estudiante, puede contribuir en la mejora de ese proceso, al planificar diseños instruccionales que atiendan a sus particularidades.

Los estilos de aprendizaje, pueden dar cuenta de las características propias de los estudiantes a la hora de aprender (Ventura 2011). Sumado a ello y en atención a las diferencias que los estudiantes presentan, la personalización del aprendizaje, se encuentra dentro de los planteos pedagógicos que proponen el ajuste de las actividades de enseñanza y de aprendizaje a las características de los alumnos, y es una de las nociones más exploradas en la actualidad en el contexto educativo. (Coll, 2016).

En esta dirección, los avances en el uso de las tecnologías de la información aplicadas a la educación (TIAE), han tenido una significativa influencia en la articulación adecuada de las necesidades, intereses y bagaje de conocimientos de los alumnos. La consolidación de la personalización, especialmente en ámbitos educativos con gran número de estudiantes, puede verse significativamente beneficiada con el apoyo tecnológico, y, por supuesto, con aquellas prácticas educativas que avalen su efectividad. (Lerís López, Vea Muniesa y Velamazán Gimeno, 2015).

La aplicación de las tecnologías ha mostrado la posibilidad de adaptar el aprendizaje a las particularidades de los alumnos. En este contexto, la adaptación del aprendizaje en ambientes virtuales permite llevar adelante una formación personalizada, mediante la utilización de TIAE, y es considerada una tendencia de innovación educativa. (Fidalgo, 2016).

1.1. La Personalización y Adaptación del Aprendizaje en Ambientes Virtuales

La personalización del aprendizaje puede entenderse como la adecuación pedagógica de contenidos y diseño del ambiente virtual para satisfacer las necesidades cognitivas, estilos y preferencias de los estudiantes, entendidos como individualidades (Klašnja-Milićević,Vesin, Ivanovic, Budimac y Jain, 2017). Si bien la idea de personalizar el aprendizaje no es una novedad en el mundo académico, su demanda ha crecido en la última década, a partir del paradigma educativo centrado en el estudiante. (Lerís López, Vea Muniesa y Velamazán Gimeno, 2015). Estos enfoques, intentan fomentar un aprendizaje más efectivo, activo, eficiente y satisfactorio para sus destinatarios. (Brusilovsky, 1996).

La idea común de un enfoque personalizado es brindar una experiencia de aprendizaje, valorando a cada individuo sea en clases presenciales, o mediante el empleo de un software educativo. Se orienta a que el estudiante recorra su propio camino hacia el conocimiento y puede ser aplicada a los contenidos, a los métodos de enseñanza y al ritmo del aprendizaje. En el ámbito del e-learning, los sistemas de gestión del aprendizaje (LMS) se utilizan para la creación de AVAs, y ofrecen una multiplicidad de herramientas que facilitan el diseño y construcción de propuestas adaptativas (Leris y Sein-Echaluce, 2011). En estos ambientes, los profesores planifican, orientan y guían el proceso de aprendizaje, generando interacciones entre estudiantes, docentes y recursos pedagógicos.

Estrechamente vinculada a la personalización, la adaptación ha planteado nuevos desafíos en el e-learning. Puede entenderse como un método para crear una experiencia de aprendizaje, que tiene en cuenta por un lado, la adaptación de los contenidos: aquellos que resulten más significativos de acuerdo a las necesidades del estudiante; y por otro, la adaptación de la presentación de dichos contenidos, para que el proceso de aprendizaje resulte más eficiente. En este caso, y dependiendo de las interacciones estudiante/AVA, el proceso se adapta al estudiante al modificar la presentación y características del material y actividades en el diseño instruccional, como respuesta a las acciones del alumno.

Los AVAs, entonces, deben considerar la diversidad en cuanto a preferencias de aprendizaje: diferentes niveles de conocimiento, habilidades cognitivas, estilos

de aprendizajes, emociones, reacciones, errores frecuentes, entre otros. Un LMS que intente poseer rasgos adaptativos debería considerar cómo identificar dicho conjunto de características (por ejemplo, mediante el uso de restricciones de acceso), a fin de atender a las necesidades y preferencias de los alumnos, entregar el material de aprendizaje adecuado y, además, proveer retroalimentación efectiva. (Klašnja-Milićević et al, 2017).

Para abordar la heterogeneidad de los perfiles cognitivos e implementar la personalización y adaptación en AVAs, en proyectos anteriores, hemos considerado los estilos de aprendizaje predominantes y el conocimiento previo de los estudiantes. Siguiendo esta línea, esta propuesta avanza en la ampliación y profundización de los procesos de personalización y adaptación, mediante el diseño de itinerarios de aprendizaje personalizados. Para ello, se identificarán los estilos de aprendizaje, conocimiento previo, y errores frecuentes detectados por los docentes o la implementación de un software específico desarrollado por nuestro grupo de trabajo. Se espera así, una mejor adecuación de la retroalimentación y presentación de materiales de estudio, hacia el logro de un mejor rendimiento en el aprendizaje.

1.2. Estilos de Aprendizaje

El concepto de estilos de aprendizaje implica, desde el punto de vista cognitivo, que los estudiantes procesan la información de diferentes maneras y que poseen diversas preferencias frente a la presentación de materiales de estudio (Moser y Zumbach, 2018). Conocer los estilos de aprendizaje puede facilitar los procesos para adquirir, interpretar y analizar la información, y, de la misma manera, la adecuación de las prácticas pedagógicas de acuerdo a los estilos de los alumnos, contribuiría a facilitar un aprendizaje significativo.

La caracterización de los diferentes estilos es un tema profusamente estudiado en el ámbito del aprendizaje, (Dunn & Dunn, 1992; Kolb, 1985; Felder & Silverman; 1988). Moser y Zumbach (2018) indican que la temática continúa siendo un tópico de discusión que genera controversias. Sin embargo se han presentado pruebas empíricas acerca de la flexibilidad y dinamismo de los estilos, característica que permitiría a los estudiantes adecuar sus estilos a los contextos y exigencias de aprendizaje. Es éste uno de los motivos que sostiene los esfuerzos que implican la personalización y adaptación del aprendizaje en un AVA, para adecuar el ambiente a las particularidades de cada estudiante, entendido como una individualidad.

1.2.1. Estilos de Aprendizaje en AVAs

El ámbito de la educación mediada por computadora ha enfatizado la utilización del conocimiento sobre estilos de aprendizaje, para personalizar y facilitar el aprendizaje. Aunque la temática es un terreno de polémicas; puede decirse al respecto que, mayoritariamente, se considera relevante que un sistema de aprendizaje adaptativo considere el estilo de aprendizaje específico de un alumno determinado (Coffield, Moseley, Hall & Ecclestone, 2004); (Popescu, 2010). La relevancia de la investigación sobre los estilos de aprendizaje en AVAs, amerita la continuidad de la indagación científica. Nuestros resultados hasta el momento, apoyan la vertiente que indica que su identificación, como parte del modelo cognitivo del estudiante, contribuye a la personalización y adaptación del aprendizaje en E-Learning, y mejora el proceso y el rendimiento del estudiante.

1.2.2. El Modelo de Felder y Silverman

Entre la multiplicidad de modelos de estilos de aprendizaje, seleccionamos el denominado modelo de Felder y Silverman (1988). Es uno de los modelos más utilizados dado que se dirige de manera específica al ámbito universitario y fue desarrollado originalmente para estudiantes de ingeniería, aunque ha sido aplicado a estudiantes de otras disciplinas. Es notorio su impacto en Latinoamérica, particularmente en Argentina y México

(Ventura, 2011). El enfoque indica que los estudiantes poseen modos particulares de seleccionar, procesar, absorber y retener información nueva, conformando patrones individuales de aprendizaje, flexiblemente estables, y que varían de persona en persona.
El instrumento diseñado por los autores es el Index of Learning Styles (ILS) (Soloman y Felder, 2005). El ILS es un cuestionario de 44 preguntas de opción múltiple que evalúa las preferencias de estilos de aprendizaje a través de cuatro dimensiones bivalentes: procesamiento de información (activo / reflexivo); percepción de información (sensitivo / intuitivo); recepción de información (verbal / visual); y comprensión de la información, (secuencial / global). Las cuatro dimensiones, varían en un rango de 1 a 11, y los resultados de las preferencias pueden ser ubicados en una posición balanceada (puntaje de 1 a 3); moderada (puntaje 5 a 7), y la preferencia predominante (puntaje 9 a 11).

1.3. Propuesta de Personalización del aprendizaje y adaptación en un Ambiente Virtual

En base a lo expuesto, nuestra propuesta se dirige ampliar y profundizar en la personalización del aprendizaje de estudiantes universitarios, mediante las posibilidades adaptativas que ofrece el LMS Moodle. Se implementará la experiencia en estudiantes de Ingeniería y de Psicología. Se realizará como complemento de las clases presenciales, y la utilización del AVA por parte de los estudiantes será voluntaria.
En los proyectos anteriores, la tarea se centró en los estilos de aprendizaje predominantes. En este proyecto ampliaremos la personalización con adaptación a todos los estilos, diseñando de esta manera estrategias de enseñanza que contemplen la diversidad de características del perfil cognitivo de los estudiantes. Para ello, se utilizarán las diferentes restricciones de acceso que ofrece Moodle.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

- **Personalización y adaptación del aprendizaje**: se estudia y explora alternativas/ estrategias de personalización y adaptación del diseño pedagógico en AVAs como complemento de clases presenciales, y en contexto de masividad de matrícula.
- **Identificación de estilos de aprendizaje:** indagación de los diferentes estilos como aportes al diseño y desarrollo de propuestas formativas personalizadas.
- **Itinerarios en AVA:** definición y desarrollo de itinerarios de aprendizaje personalizado, con materiales de estudio y actividades con retroalimentación adecuada a estilos de aprendizaje, niveles de conocimiento y errores frecuentes detectados en el dominio de estudio.

3. RESULTADOS Y OBJETIVOS

- Evaluar el conocimiento previo usando la aplicación InfoSem[1] (Huapaya, Lizarralde, y Vivas, 2015).
- Detectar errores frecuentes mediante la aplicación GEMP[2].
- Definir segmentos de estudiantes, por nivel de conocimiento previo y estilos de aprendizaje.
- Desarrollar materiales de estudio y pruebas personalizadas.
- Investigar la adaptación dinámica con herramientas y recursos de Moodle.
- Explorar la interacción del estudiante en el AVA, con respecto a la búsqueda y selección de materiales instruccionales en relación con los estilos de aprendizaje.

[1] INFOSEM: sistema computacional de evaluación cognitiva del conocimiento, desarrollado para la construcción de redes semánticas.

[2] GEMP: Gestor de Evaluaciones utilizando el Modelo de Perturbación

4. FORMACIÓN DE RECURSOS HUMANOS

En el equipo de investigación hay especialistas en educación, informática, ingeniería y psicólogos educativos. Asimismo, los integrantes del equipo de auxiliares de práctica de la materia Fundamentos de la Informática, colaboran con el desarrollo e implementación de los ejercicios de autoevaluación y selección de materiales de estudio.

5- BIBLIOGRAFIA

Brusilovsky, P. (1996). Methods and techniques of adaptive hypermedia. User Modeling and User Adapted Interaction, Vol. 6, no. 2-3, pp 87-129.

Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). Learning styles and pedagogy in post-16 learning: A systematic and critical review. Learning and Skills Research

Coll, C. (2016). La personalización del aprendizaje escolar. El qué, el por qué y el cómo de un reto insoslayable. En J. M. Vilalta (Dr.). Reptes de l'educació a Catalunya.Anuari d'Educació 2015 Barcelona: Fundació Jaume Bofill. Traducción de Iris Merino.

Dunn, R., & Dunn, K. (1992). Teaching elementary student through their individual learning styles. Boston: Allyn & Bacon.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching in engineering education. Engineering Education, 78(7), pp. 674–681.

Huapaya, R; Lizarralde, F. y Vivas, J. (2015). Modelo para visualizar y evaluar el conocimiento conceptual. Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación nº 15, pp. 14-24. Red UNCI. Universidad Nacional de La Plata.

Klašnja-Milićević, A; Vesin, B; Ivanovic, M; Budimac, Z. y Jain, L. (2017). E-Learning Systems Intelligent Techniques for Personalization. Springer. 2017. (eBook) Part 2 Chapter 2 pp. 21-26

Kolb, D. (1985). Learning style inventory. Boston: McBer & Co.

Leris, D. y Sein-Echaluce, M.L. (2011). Aprendizaje: un objetivo del paradigma educativo centrado en el aprendizaje. ARBOR Ciencia, Pensamiento y Cultura, Vol. 187, pp 123-134, Extra 3.

Lerís López, D; Vea Muniesa, F.y Velamazán Gimeno, A. (2015). Aprendizaje adaptativo en moodle: tres casos prácticos. E K S vol.16 nº 4. pp. 138-157.

Moser, S. y Zumbach, J., (2018). Exploring the development and impact of learning styles: An empirical investigation based on explicit and implicit measures. Computers & Education 125, pp.146–157.

Popescu, E. (2010). Adaptation provisioning with respect to learning styles in a w-based educational system: an experimental study. Journal of Computer Assisted Learning, 26(4), pp. 243–257.

Soloman, B. A. y Felder, R. (2005). Index of learning styles questionnaire. NC State University.. Disponible en: https://www.webtools.ncsu.edu/learningstyles/

PROYECTO DE LABORATORIOS REMOTOS PARA LA ENSEÑANZA DE LA FÍSICA.

María Joselevich[*1,2,7], Roberto E. Alonso[1,3,4,7], Pablo González Casco[1,7], Martín Morales[1,5,7], Jorge Osio[1,7], Alejandra Serial[2,6,7]

[1]Instituto de Ingeniería y Agronomía, Universidad Nacional Arturo Jauretche (UNAJ), Argentina.

[2]Centro de Política Educativa,UNAJ, Argentina.

[3]Departamento de Ciencias Básicas , Facultad de Ingeniería, Universidad Nacional de La Plata, Argentina

[4]Instituto de Física La Plata-CONICET, UNLP, La Plata, Argentina

[5]Universidad Tecnológica Nacional - Facultad Regional La Plata, Argentina

[6]Instituto de Ciencias Sociales y Administración, UNAJ, Argentina.

[7]Programa de Tecnologías de la información y la comunicación (TIC) en aplicaciones de interés social, IIyA, UNAJ

{mjoselevich, ralonso, pcasco, martin.morales, josio, aserial}@unaj.edu.ar

Resumen

Los laboratorios remotos son espacios físicos reales que contienen dispositivos experimentales a los que se puede acceder y manipular desde sitios remotos a través de internet. Se utilizan para realizar prácticas de laboratorio en tiempo real. En este trabajo se presenta el proyecto de diseño e implementación de un laboratorio remoto de física para el abordaje de las ondas mecánicas en las etapas iniciales de los estudios universitarios en carreras de ingeniería.

Palabras clave: *Laboratorio Remoto, Educación a distancia, Tecnología informática aplicada a Educación*

Contexto

Este trabajo está enmarcado en el proyecto de investigación "*Uso de laboratorios remotos en la enseñanza de la física: estudio de una propuesta didáctica",* actualmente en ejecución en la Universidad Nacional Arturo Jauretche. Integra las propuestas en desarrollo en el *Área de Herramientas TIC para el Diseño de Estrategias Didácticas Educativas* del *Programa de Investigación Tecnologías de la Información y la Comunicación (TIC) en Aplicaciones de Interés Social* sito en el Instituto de Ingeniería y Agronomía (UNAJ) y se desarrolla en el marco de un convenio de colaboración institucional con el Departamento de Física de la Facultad de Ciencias Exactas y Naturales (UBA). El equipo que lleva adelante el proyecto está integrado por docentes de los Institutos de Ciencias Sociales y Administración (ICSyA) y de Ingeniería y Agronomía (IIyA), e investigadoras del Centro de Política Educativa (CPE).

1. Introducción

Desde el punto de vista de la didáctica específica de las ciencias naturales, la utilización adecuada de recursos informáticos parece ser potenciadora y enriquecedora de los procesos de aprendizaje (Aksela, 2005). Los estudiantes tienen acceso a una enorme cantidad de información de fuentes impensadas hasta hace muy poco tiempo. Pueden procesar datos y mediciones hechas por ellos mismos con herramientas de creciente precisión o por investigadores que viven en cualquier parte del planeta y utilizan equipos altamente costosos y sofisticados. Este contexto los estimula a ser sujetos sumamente

activos de su propio aprendizaje, dejarse guiar por sus propios intereses e inquietudes y por las preguntas que les surgen legítimamente. Esta manera de transitar el aprendizaje es sumamente distinta a la transmisión tradicional del conocimiento y brinda a los equipos docentes oportunidades para potenciar enormemente su rol de guía y orientador de los aprendizajes del alumnado

Entre las ciencias, la física es por excelencia una ciencia natural por lo que la experiencia siempre resultó la motivación, el objeto de la modelización y de su formulación matemática. La comprensión de los fenómenos naturales en su complejidad desde los inicios involucró a la experiencia en una versión "simplificada" en el laboratorio, donde el investigador pretende tener control individual sobre las variables relevantes del problema en cuestión, algunas veces con bastante éxito. A partir de las experiencias en el laboratorio y el control de variables es de donde surgen algunos modelos con los que se tratan de describir y comprender los fenómenos por lo que resultan parte vital en la enseñanza. A partir de las situaciones recreadas en el laboratorio es posible para los estudiantes realizar conexiones entre el fenómeno y el modelo, además de desarrollar la capacidad crítica, el libre pensar, refutar y crear, dado que allí se enfrentan con las mediciones sobre sistemas "reales". Entre los avances en la enseñanza relacionados con las Tecnologías de la información y la comunicación (TIC), el uso de sensores o colectores electrónicos de datos impacta fuertemente en el desarrollo experimental de la enseñanza de las ciencias naturales. El trabajo con sensores permite tomar y procesar una gran cantidad de información de procesos que pueden ocurrir en tiempos muy cortos o en tiempos muy largos, con un alto nivel de detalle y una gran facilidad en la manipulación tanto de los equipos como de los datos. Según numerosos estudios (Tortosa, 2012), el uso de sensores enlazado a computadoras, si se da en ambientes de aprendizaje adecuados, conlleva ventajas en el desarrollo de habilidades cognitivas de orden superior.

Si bien los recursos presentados más arriba son señalados como facilitadores del aprendizaje, los contextos universitarios actuales hacen que existan limitaciones en cuanto a tiempo, equipos y personal por las cuales es extremadamente difícil abordar todos y cada uno de los conceptos discutidos en un curso, en una práctica de laboratorio (Gil, 1997). En este sentido, el avance en el desarrollo de herramientas TIC, permite pensar en diferentes estrategias para enfrentar estas limitaciones. Este es el caso de los laboratorios remotos (LR) como los que han incorporado numerosas universidades del mundo, entre las que se encuentran Stanford, Universidad de Australia del Sur, Deusto y, en la Argentina, la Universidad Nacional del Litoral y la Universidad Nacional de Rosario.

Los LR son espacios físicos reales donde se encuentran equipos manipulables a través de internet. Contienen dispositivos experimentales con el agregado de sensores, cámaras y mecanismos de automatización y control externo para que los usuarios puedan acceder a la manipulación, registro y controles desde sitios remotos en tiempo real.

Desde el punto de vista de la enseñanza, los LR tienen dos ventajas fundamentales. Por un lado, permiten acrecentar la cantidad de tiempo y la frecuencia de acceso de los estudiantes al equipamiento de laboratorio respecto a las propuestas presenciales. Por otra parte, la disponibilidad de nuevas herramientas didácticas invita a los docentes a enriquecer e innovar en sus prácticas pedagógicas.

Además de lo anterior, el uso de LR permite el ahorro de recursos económicos, la extensión de recursos materiales escasos y el compartir equipamiento con otras instituciones a la vez que proporciona acceso a equipos que pueden ser demasiado costosos, peligrosos o logísticamente problemáticos para una institución de enseñanza.

Se han publicado varias experiencias educativas en el uso de LR para la enseñanza de la física (C. Arguedas Matarrita y S. B. Concari, 2016). Entre ellas se puede

mencionar la investigación en ámbitos universitarios de Marques (Marques y otros, 2014) y, en el caso de la Argentina los trabajos de los grupos ya mencionados de la Universidad Nacional del Litoral y de la Universidad Nacional de Rosario (Kofman y Concari, 2011 y Lerro y Marchisio, 2016). Diferentes investigadores se han dedicado a investigar distintos aspectos relacionados con el trabajo con los LR, que abarcan desde el establecimiento de criterios de idoneidad didáctica de los desarrollos (Arguedas–Matarrita y otros, 2017) hasta distintos aspectos del abordaje del trabajo de los estudiantes en los LR (Lerro, F. y Marchisio, S., 2016).

Descripción de la experiencia

La Universidad Nacional Arturo Jauretche, siguiendo los lineamientos de su Proyecto Institucional, diseña e implementa distintas estrategias de enseñanza que, atendiendo a la heterogeneidad de sus estudiantes, buscan producir algunas transformaciones en relación con el modelo educativo universitario "tradicional" y lograr mayores niveles de retención y graduación manteniendo la excelencia académica y los objetivos de superación y formación humana y profesional.

En ese marco, se están desarrollando distintas propuestas de trabajo sistemático en equipos interdisciplinarios como el que propone este proyecto, el cual se enmarca en las acciones dirigidas al trabajo en la asignatura Física 1, correspondiente al ciclo de las ciencias básicas de las seis carreras de Ingeniería que se ofertan en la UNAJ.

Los objetivos de esta asignatura incluyen la formación de los estudiantes en el desarrollo de ciertas competencias que se consideran fundamentales para sus estudios posteriores, entre ellas, aquellas relacionadas con el pensamiento y trabajo experimental. Este trabajo requiere, para llevarse a cabo, del uso de laboratorios e instrumental específico que no se encuentra disponible dada la insuficiencia infraestructura disponible para la gran cantidad de estudiantes de los primeros años.

Lo anterior motivó a este grupo de trabajo a explorar estrategias de enseñanza que contribuyan a la formación práctica de los y las ingenieros/as, pensadas para auxiliar en la circunstancia actual y, eventualmente, complementar las actividades que se realicen en el futuro.

Entre las estrategias posibles, se eligió estudiar la incidencia de la incorporación de laboratorios remotos (LR) en las prácticas de enseñanza de la materia Física 1. El trabajo que se presenta se dirige en particular al estudio de las ondas mecánicas tomando en cuenta los antecedentes del tratamiento de este tema a partir del análisis de ondas sonoras (Bocanegra Cifuentes, J., 2012), ondas en una cuerda (González, G., 2005), entre otras, utilizando distintos recursos tecnológicos como laboratorios virtuales, tutoriales en videos, blogs. Los laboratorios remotos, se presentan actualmente como una opción que despierta un gran interés.

La implementación del laboratorio remoto consta de las siguientes etapas:

- El servidor de enlace fue implementado mediante la plataforma educativa moodle, en donde se configuró un plugin para la solicitud de turnos y la gestión de acceso al laboratorio.
- El servidor multimedia se implementó en una mini computadora (raspberry pi), donde se configuró Motion como un servidor streaming para el acceso a las imágenes y Apache como servidor web en donde se muestran las opciones de encendido y configuración del laboratorio.
- Se utilizó el motor de base de datos *phpMyAdmin* que permite manejar la administración de MySQL para el almacenamiento de los datos generados en cada experiencia.

- Almacenamiento de los datos en la nube mediante los servicios del proveedor google.
- Para el control a distancia del instrumental se creó una API rest con un conjunto de comandos que permiten seleccionar las diferentes opciones de configuración

2. Líneas de investigación y desarrollo

Aspectos pedagógico didácticos

Considerando que el equipo de investigación que desarrolla este proyecto incluye docentes de las asignaturas para las cuales se diseñará e implementará la propuesta didáctica que se utilizará en este estudio, el abordaje metodológico que se propone se enmarca en la línea de investigación-acción.

Se entrevistará a los estudiantes a fin de indagar acerca de los usos y los sentidos de sus interacciones con el laboratorio remoto a lo largo del proceso de implementación de la propuesta didáctica. Asimismo, se realizarán observaciones de clases y se entrevistará a los integrantes del equipo docente, incluyendo a quienes participan del diseño e implementación de la propuesta didáctica.

Aspectos del diseño e implementación del sistema

La implementación del sistema requiere de la integración de un conjunto de herramientas informáticas que posibiliten el acceso y control de forma remota del laboratorio **(**Arguedas-Matarrita y Concari, 2016**).** Para su implementación se investigaron algunos aspectos (Bocanegra Cifuentes, 2012), que permitieron la correcta implementación y visualización del laboratorio en tiempo real :

- Servidor de enlace: Para Gestión de acceso a usuarios
- Servidor multimedia: Provee imágenes y audio del ambiente experimental
- Servidor de laboratorio: Da soporte a la aplicación y gestiona la base de datos
- Seguridad para garantizar el acceso al manejo del instrumental solo a personas autorizadas

3. Resultados Obtenidos / Esperados

Actualmente el proyecto se encuentra en la etapa de testing, en donde se ha logrado la implementación completa del sistema con algunas pruebas de funcionamiento que permitieron verificar la correcta ejecución del acceso de usuarios, solicitud de turnos, acceso a la plataforma donde se pueden ejecutar los comandos de configuración de forma remota (selección de frecuencia, fase y obtención de imágenes del experimento) y el almacenamiento de datos del experimento en la nube.

La siguiente instancia consiste en resolver los aspectos de seguridad para poder realizar la primer prueba piloto con una comisión de alumnos, en donde se podrán recabar datos sobre las dificultades surgidas, los aportes del mismo al tema de estudio y las posibles modificaciones a futuro.

El presente proyecto de investigación supone que el uso de LR en el diseño y el desarrollo de actividades didácticas promoverá el enriquecimiento de las prácticas de enseñanza en los docentes. Se espera además que la disponibilidad de LR en el marco de la propuesta de enseñanza que se diseñará fortalezca la autonomía de los estudiantes respecto a la apropiación del conocimiento.

Esperamos que estos resultados impacten en los estudiantes de manera de aumentar sus aprendizajes acerca de los contenidos que se abordan.

4. Formación de recursos humanos

En el marco de este proyecto, se han desarrollado las Prácticas Profesionales

Supervisadas (PPS) de dos estudiantes de la carrera de Ingeniería en Informática de UNAJ. Se prevé que la participación de estudiantes continúe por un lado en el marco del desarrollo de una beca ya adjudicada por el Consejo Interuniversitario Nacional y por otro con una nueva PPS.

5. Bibliografía

Aksela, M. (2005). Supporting meaningful chemistry learning and higher-order thinking through computer-assisted inquiry: A design research approach.

C. Arguedas-Matarrita y S. B. Concari (2016) Laboratorios remotos para la enseñanza de la física: características tecnológicas y pedagógicas Revista de Enseñanza de la Física. Vol. 28, No. Extra, Nov. 2016, 235-243

C. Arguedas–Matarrita, S. B. Concari y B. Giacomone (2017) La idoneidad didáctica de los laboratorios remotos como recursos para la enseñanza y aprendizaje de la física. Revista de Enseñanza de la Física. Vol. 29, No. Extra, Nov. 2017, 511–517

Bocanegra Cifuentes, J. (2012). Laboratorio de ondas mecánicas apoyado con Tecnologías de la Información y la Comunicación de libre acceso. Tesis de Maestría, Universidad Nacional de Colombia.

Gil, S., (1997). Nuevas tecnologías en la enseñanza de la física oportunidades y desafíos. Educación en Ciencias, 1(34): p. 1-10.

González, G. (2005). Enseñanza de ondas mecánicas en cuerdas usando Matemática. Universitas Scientarium. Vol 10. Número especial. Enero-Junio. Bogotá, Colombia. pp17-24.

Kofman, H. y Concari, S. (2011). Using remote labs for Physcis teaching. En: Using Remote Labs in Education. J. García Zúbia and G. Alves (Eds.): University of Deusto Publications, 293-308. Recuperado de http://www.deusto-publicaciones.es/deusto/pdfs/otraspub/otraspub01.pdf (13/03/2019).

Lerro, F. y Marchisio, S. (2016). Preferences and uses of a remote lab from the students' viewpoint. International Journal of Online Engineering, 12(3), 53-57. Recuperado de http://onlinejournals.org/index.php/i-joe/article/view/5468/3855 (13/03/2019).

Marques, M. A., Viegas, M. C., Costa-Lobo, M. C., Fidalgo, A. V., Alves, G. R., Rocha, J. S. y Gustavsson, I. (2014). How Remote Labs Impact on Course Outcomes: Various Practices Using VISIR. IEEE Transactions on Education, 57(3), August 2014. Recuperado de http://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6645452 (13/03/2019).

Tortosa, M. (2012). The use of microcomputer based laboratories in chemistry secondary education: Present state of the art and ideas for research-based practice. Perspective. Chem. Educ. Res. Pract., 13, pp. 161-171.

Proyecto IDEAS: el caso de un entorno virtual de enseñanza y aprendizaje

Sanz Cecilia[1,2], Iglesias Luciano[1], Salazar Mesía Natalí[1,3], Archuby Federico[1, 3], Alejandra Zangara[1], Buffarini Abril[1]

[1] Instituto de Investigación en Informática LIDI (III-LIDI). Centro Asociado CIC
Facultad de Informática – Universidad Nacional de La Plata
[2]Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires
[3]Becario UNLP
{csanz, nsalazar, farchuby}@lidi.info.unlp.edu.ar, li@info.unlp.edu.ar, Alejandrazangara@gmail.com, abril.buffarini@gmail.com

RESUMEN

En este trabajo se presenta un proyecto de investigación y desarrollo enfocado en los entornos virtuales de enseñanza y aprendizaje. Se enmarca en un proyecto más amplio del Instituto de Investigación en Informática LIDI, en el que se estudian metodologías y herramientas de las Ciencias de la Computación para realizar aportes en escenarios educativos. En el marco del proyecto IDEAS se estudian tecnologías, funcionalidades y aspectos de diseño de sistemas centrados en la web cuyo objetivo es la mediación de cursos con diferentes modalidades educativas. Como parte del proyecto se ha implementado el entorno IDEAS que se utiliza en el marco de la Facultad de Informática de la UNLP. Además se cuenta con un sitio web para difusión del proyecto y compartirlo con la comunidad.
Se presentan en este trabajo los principales resultados alcanzados en relación al proyecto.

Palabras claves: entornos virtuales, entornos virtuales de enseñanza y aprendizaje, sistemas web.

CONTEXTO

Esta investigación forma parte del 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" del III-LIDI, Facultad de Informática, Universidad Nacional de La Plata, y acreditado por el Ministerio de Educación de la Nación. Al mismo tiempo, se trata de un proyecto que ha sido apoyado por la Facultad de Informática de la UNLP.

1. INTRODUCCIÓN

Desde hace ya varios años las universidades y las instituciones educativas en general, han visto el crecimiento de propuestas educativas en las que se han integrado diferentes tecnologías digitales con el fin de enriquecer, y ampliar las estrategias didácticas y metodológicas utilizadas. De esta manera se han generado propuestas en las que se extiende el aula, más allá del espacio físico en el que se desarrollan los encuentros cara a cara o las tradicionales clases. En algunos casos las propuestas se centran en la mediación de gran parte de las actividades a través de entornos digitales (Sanz y Zangara, 2012; Zangara, 2014; Sanz, 2015).
Los entornos virtuales de enseñanza y aprendizaje (EVEA) han jugado un papel fundamental en estas tendencias. Se trata de sistemas centrados en la web, que se orientan a la gestión y diseño de cursos y facilitan la mediación de procesos educativos a través de las múltiples herramientas que ofrecen (Castañedo Quinteros y López Vicent, 2007; Jaggars y Xie, 2016). Entre sus principales herramientas se mencionan: las orientadas a la interacción entre docentes y alumnos y alumnos entre sí, herramientas para la publicación de la propuesta y de los contenidos y actividades, para la colaboración, para la evaluación y para el seguimiento de los estudiantes. Sumados a estos EVEA, se ha avanzado también sobre el diseño y desarrollo de Campus Virtuales, otro tipo de sistemas que integran servicios que forman parte de las instituciones educativas, como puede ser servicios de biblioteca, espacios para la interacción y socialización de toda la comunidad educativa, repositorios de contenidos, entre otros. Los campus suelen además integrar las

funcionalidades de los EVEA (Castañedo Quinteros y López Vicent, 2007).

Los EVEA han ido evolucionando con el correr de los años integrando más funcionalidades. En la actualidad se debate la percepción de estos sistemas por parte de los alumnos que suelen considerarse como "visitantes" de estos espacios en lugar de "residentes", acorde a los estudios de Dabbagh y Fake (2017). La inclusión de herramientas sociales puede aportar a que los estudiantes se apropien de mejor manera de estos EVEA. Por otra parte, técnicas tales como las del área de Computación Afectiva y minería de texto se vienen utilizando para dotar a estos sistemas de nuevas capacidades para orientar a los alumnos, ayudar en su autorregulación y en su aprendizaje en general.

Otro aspecto estudiado en el proyecto se enfoca en el diseño de los cursos en los EVEA. En este sentido, se ha tomado como base la Teoría de Distancia Transaccional de Moore (1972, 1989, 2012) y sus componentes de diálogo, estructura y autonomía.

En este proyecto se estudian estas temáticas, y se avanzan en el diseño e implementación de un EVEA que se denomina IDEAS y es utilizado en el marco de la Facultad de Informática de la UNLP.

2. LINEAS DE INVESTIGACION / DESARROLLO

Las principales líneas de investigación y desarrollo en vinculación a este proyecto son:

- Diferenciación del concepto de EVEA de otros sistemas web educativos
- Funcionalidades y arquitecturas de los EVEA
- Computación afectiva en EVEA
- Indicadores y herramientas para el seguimiento de estudiantes y de procesos grupales y/o colaborativos en EVEA. Trazabilidad
- Comunicación e interacción en EVEA
- Estándares de publicación de contenidos
- Componentes del diseño de un curso mediado en un EVEA

3. RESULTADOS OBTENIDOS

Se presentan aquí los principales resultados que se han alcanzado en el proyecto. En el marco de la investigación se han estudiado antecedentes sobre técnicas de análisis de sentimiento sobre textos/aportes en foros y mensajes en los EVEA (Rodriguez, Baldasarri, Sanz, 2017). Se han analizado diferentes EVEA y comparado sus funcionalidades, considerando criterios tecnológicos y pedagógicos (Sanz, 2018). Desde 2015 se ha realizado una reingeniería del entorno virtual de enseñanza y aprendizaje WebUNLP que ha evolucionado en un nuevo sistema llamado IDEAS.

IDEAS es un entorno virtual de enseñanza y aprendizaje que permite la gestión y el diseño de cursos (Figura 1). Se ha hecho un esfuerzo para realizar un sistema *responsive* y se ha trabajado en aspectos relacionados con su accesibilidad. En cada curso intervienen diferentes roles como son los de docente y alumno. A su vez los docentes pueden ser: tutores, profesores, auxiliares, entre otros. Cada uno con diferentes permisos sobre el curso. Las herramientas del curso se organizan en áreas, que tienen una intención pedagógica diferente, entre las cuales se encuentran: Inicio, Comunicación, Itinerario, Colaboración, Evaluación y Otros recursos. Las áreas se relacionan con mayor énfasis con alguno de los componentes de la Teoría transaccional de Moore (1972), como se observa en la Figura 2.

Los profesores del curso también pueden contar con un área de Gestión de alumnos y de docentes y otra de Gestión estadística. En esta última se han desarrollado diferentes formas de visualizar información de interés sobre las actividades del curso (ver Figura 3). Por ejemplo, se puede ver para cada contenido publicado la cantidad de accesos, o para la mensajería cuántos mensajes enviaron los alumnos a sus docentes y viceversa. Su objetivo es aportar a la trazabilidad y seguimiento de alumnos, y a la toma de decisiones para la intervención didáctica.

Por otra parte, en el trabajo dentro del curso es posible siempre tener las novedades y se pueden recibir notificaciones en el correo electrónico de cada participante, acorde a su configuración (notificaciones semanales, diarias o por eventos para cada herramienta). Esto fue pensado para fomentar que los participantes de un curso puedan estar atentos a las interacciones y novedades que se producen.

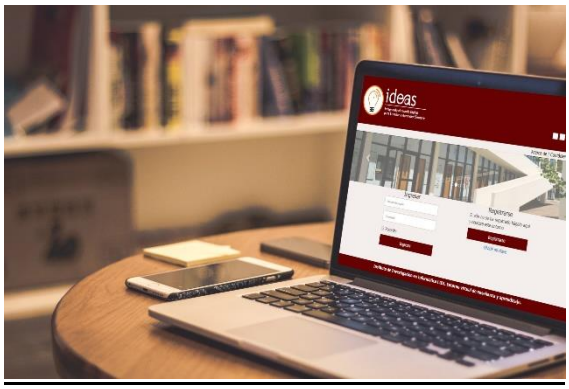


Figura1. Pantalla principal de IDEAS

Área	Herramientas	Relacionado con	Comentarios
Inicio	Bienvenida	Estructura	Identifica las características del curso y presenta su estructura. Informa sobre canales adicionales de comunicación
Itinerario	Propuesta Materiales y actividades	Estructura	Organizador de la estructura de la propuesta. Se presenta la estructura en dos niveles: CURSO: Programa, cronograma MATERIALES: contenidos presentados en diferentes lenguajes, según decisiones de transposición didáctica y mediación.
Comunicación	Mensajería Cartelera Foro	Diálogo	Columna vertebral del diálogo del curso.
Colaboración	Repositorios Grupos	Estructura	Información de los asistentes al curso Repositorio de materiales de trabajo colaborativo [Pueden funcionar en vinculación con las herramientas de diálogo]
Evaluación	Autoevaluación Tareas	Estructura Diálogo	Organización de las instancias de evaluación. En las actividades menos estructuradas, puede haber atención del tutor o docente.
Otros recursos	Medioteca Glosario	Estructura	Repositorios de materiales y presentación de información en el glosario

Figura 2 – Áreas y herramientas del curso en IDEAS y su relación con los componentes de diálogo y estructura de la Teoría de Distancia Transaccional.

Figura 3 – Visualizaciones de información estadística de seguimiento del curso. Situación de los participantes respecto de la entrega de una tarea.

Un aspecto destacable de IDEAS ha sido el área de comunicación para atender al componente de diálogo de una propuesta educativa mediada por este sistema. Se ha puesto especial énfasis en el diseño de una mensajería interna al curso que acompaña la tarea de los tutores y docentes en general. Constituye una de las herramientas centrales del EVEA para acompañar el diálogo entre los participantes. Por otra parte, en los foros se ha diseñado la posibilidad de visualizar los hilos de comunicación que se generan (ver Figura 4), y se está trabajando sobre indicadores de seguimiento sobre esta herramienta para mostrar en el área de Gestión Estadística (Sanz, Zangara & Dieser, 2016).

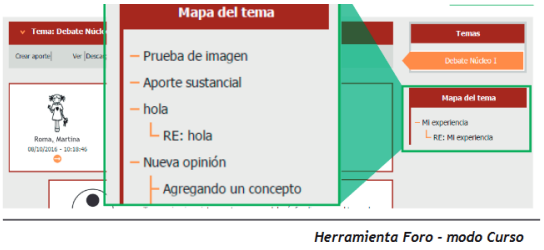


Figura 4. Herramienta de Foro con el mapa del tema que se desagrega en sus hilos de comunicación.

Otros aspectos que se vienen trabajando para integrar a IDEAS son:

- Estrategias lúdicas: se creó un juego llamado Desafiate para dispositivos móviles que toma las evaluaciones de un curso de IDEAS y las presenta como una aventura dentro del juego (Archuby, Sanz y Pesado, 2018).
- Herramienta de comunicación sincrónica para reuniones virtuales: se ha creado, a partir de una tesina de grado, una herramienta llamada InfoMeeting (Digiani, Gorga y Sanz, 2018) que actualmente se está integrando a IDEAS.
- Herramientas de Realidad Aumentada: se espera avanzar en la integración de actividades con Realidad Aumentada dentro de un curso en IDEAS.
- Comunidad IDEAS: se ha creado un sitio web llamado Proyecto IDEAS (https://proyectoideas.info.unlp.edu.ar/) donde se encuentra documentación del sistema, videotutoriales para la comunidad interesada y la historia del proyecto. La propuesta es en un futuro próximo compartir el desarrollo.

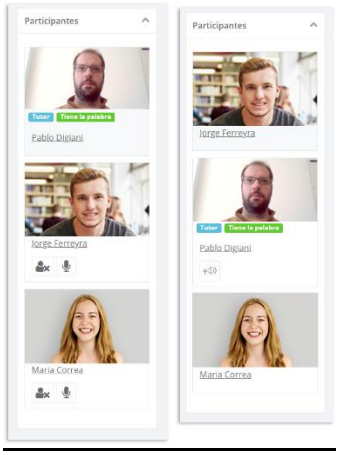


Figura 5. Pantalla de Participantes de Infomeeting

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 están realizando la tesis de maestría en TIAE.
- Una pasante del área de diseño y comunicación visual.

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; Sanz, C. y Pesado, P. (2018) Desafiate: un juego para la autoevaluación. Actas del XIII Congreso sobre Tecnología en Educación & Educación en Tecnología (TE&ET 2018), ISBN: 978-950-766-124-2, págs. 209-212, 2018. Disponible en: http://sedici.unlp.edu.ar/handle/10915/69073
- Castañeda Quintero, L. y López Vicent, P. (2007). “Entornos Virtuales de Enseñanza Aprendizaje Libres: MOODLE”. En Prendes Espinosa, M. P. Herramientas Telemáticas Para La Enseñanza Universitaria En El Marco Del Espacio Europeo De Educación Superior. Grupo de Investigación de Tecnología Educativa. Universidad de Murcia. ISBN: 978-84-611-7947-3. Disponible en: https://digitum.um.es/xmlui/bitstream/10201/13417/1/moodle.pdf
- Dabbagh, N. y Fake, H. (2017). 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
- Digiani, P.; Sanz,C. y Gorga, G. (2018) InfoMeeting: una herramienta de comunicación sincrónica moderada. Aprobada en 2018. Facultad de Informática. UNLP.
- Jaggars, S. & Xu, D. (2016). How do online course design features influence student performance? Computers & Education. Volume 95, 270-284.
- Moore, M. (1972). Learning autonomy: the second dimension of indepen-dent learning, Conference Fall, págs 76-88.USA.
- Moore, M. (Editor) (1989). Readings in Principles of Distance Education (Readings in Distance Education Series No. 1). American Center for the study of Distance Education. College of Education. Pennsylvania State University.
- Moore, M. (Editor) (2012) The Handbook of Distance Education. Third Edition. New York, Routledge.
- Rodriguez, V., Baldasarri, S., y 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.
- Sanz, C. (2015). Revista Bit & Byte. Nro 1. Pp -33-35. Facultad de Informática. Universidad Nacional de La Plata.
- Sanz, C.; Zangara, A. (2012). "La escritura colaborativa como e-actividad". En: XVIII Congresoargentino de ciencias de la computación.
- Sanz, C., Zangara, A. y 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, 2016

- Sanz, C. (2018) Presentaciones del Seminario de Educación a Distancia. Maestría en Tecnología Informática Aplicada en Educación. Facultad de Informática. UNLP
- Zangara, A. (2014) "Apostillas sobre los conceptos básicos de educación a distancia o …una brújula en el mundo de la virtualidad".
- Zangara, A. y Sanz, C. (2018) Interacción e Interactividad en el trabajo colaborativo mediado por tecnología informática. Tesis doctoral. Faculta de Informática. UNLP. Disponible en: http://sedici.unlp.edu.ar/handle/10915/67175

Recomendación automática de grupos en ambientes ubicuos

Quintana Cancinos, Fernando D. y Alvarez, Margarita M.
Instituto de Investigaciones en Informática y Sistemas de Información
Departamento de Informática - Facultad de Ciencias Exactas y Tecnologías
Universidad Nacional de Santiago del Estero (UNSE)
e-mail: f.quintana88@yahoo.com.ar, alvarez@unse.edu.ar

RESUMEN

La formación de grupos de aprendizaje eficaces representa uno de los principales factores de éxito en el Aprendizaje Colaborativo. Sin embargo, las características distintivas de los alumnos hacen que la operación de formar grupos de aprendizaje adecuados sea una tarea complicada. A todo esto, se suma que, con el surgimiento del aprendizaje ubicuo, el agrupamiento de los estudiantes en grupos de aprendizaje debe contemplar, además, la movilidad y las condiciones contextuales.

Esta problemática requiere del diseño de un método que genere un grupo de aprendizaje cuando un estudiante necesita realizar una actividad colaborativa. En esta línea de investigación se diseñará un método de formación de grupos que contemple para su formación que los compañeros del grupo no hayan realizado esa actividad, que tengan los pre-requisitos académicos aprobados y que se encuentren lo más cerca posible al estudiante.

Además, esta línea de investigación tiene por finalidad diseñar y construir el prototipo que genere recomendaciones de compañeros con los que un estudiante puede conformar un grupo de trabajo para desarrollar una tarea colaborativamente en un ambiente de aprendizaje ubicuo.

***Palabras clave*:** Aprendizaje Colaborativo, Aprendizaje Ubicuo, Recomendaciones, Formación de grupos.

CONTEXTO

Este trabajo se desarrollará en el marco de una Beca del Programa de Beca EVC-CIN, Resolución Nº: 389/18. Además, se inserta en el proyecto "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas" (Durán et al., 2017), correspondiente a la convocatoria 2017-2020 de la Secretaría de Ciencia y Tecnología de la Universidad Nacional de Santiago del Estero (SICYT - UNSE).

1. INTRODUCCIÓN

Tradicionalmente, el aprendizaje se ha configurado como un proceso mediante el cual el aprendiz entra en contacto y adquiere conocimientos o destrezas, de alguna fuente autorizada, pero en la actualidad, las teorías más avanzadas sobre el aprendizaje sostienen que los aprendices no absorben de forma pasiva el conocimiento, sino que más bien lo crean de forma activa, a partir de su experiencia del mundo.

(Burke, 2011) menciona que el aprendizaje activo comprende una gran variedad de prácticas, y destaca que el aprendizaje colaborativo (AC) es un elemento importante en la teoría y la práctica del este.

El AC representa un enfoque educativo que (Dillenbourg, 1999) define como "una situación, en la cual dos o más personas aprenden o buscan aprender algo juntas". Se desarrolla en una colectividad no competitiva, en la cual todos los miembros del grupo colaboran en la construcción del conocimiento y contribuyen al aprendizaje de todos.

Los estudios demuestran que se requieren tres condiciones clave para cualquier AC exitoso: las características de la tarea, los medios de comunicación y la composición del grupo (Dillenbourg, 2002). Es por ello, que la formación de grupos de aprendizaje eficaces representa uno de los factores importantes que determinan la eficiencia del AC (Amara & Santos, 2016).

Por otro lado, en los últimos años, tanto Internet como la tecnología inalámbrica han registrado un gran avance permitiendo el acceso, gracias al uso de dispositivos móviles,

a una gran variedad de información y servicios y, por lo tanto, extiende el aprendizaje haciendo realidad la idea de aprendizaje en cualquier lugar y en cualquier momento. Así, las interacciones pueden concretarse donde los estudiantes lo necesiten, en diferentes campos de la vida diaria, sin restricción de espacio ni de tiempo (Graf y Kinshuk, 2008).

Estos avances en las teorías educativas y en las tecnologías móviles han permitido el surgimiento del aprendizaje ubicuo (u-learning) que designa al conjunto de actividades formativas, apoyadas en tecnología, que están accesibles en cualquier lugar y desde cualquier dispositivo (Saadiah y Kamarularifin, 2010).

Este nuevo modelo de aprendizaje fomenta el aprendizaje autónomo, contextualizado y personalizado, siendo la personalización una de sus características principales. Esto implica adaptar las metodologías de aprendizaje a los modos de aprendizaje de los estudiantes y, ofrecerles recomendaciones personalizadas sobre materiales de estudio, actividades, tutores y grupos de aprendizaje. Para personalizar las aplicaciones ubicuas se consideran aspectos tales como: estilo de aprendizaje, conocimientos previos del estudiante e información del contexto (la ubicación, el momento, los diferentes dispositivos, etc.) (Won-Ik Park et al., 2010).

Por otra parte, la adecuada formación del grupo es un elemento importante que puede determinar el éxito o fracaso en el alcance de la meta. La formación de grupos colaborativos sin unas consideraciones cuidadosas (por ejemplo, generación de grupos en forma aleatoria), generalmente causan problemas como la participación desproporcionada de individuos, desmotivación y resistencia al trabajo grupal en actividades futuras (Cruz & Isotani, 2014).

Por eso, para la formación de grupos se debe considerar las diversidades sociales, culturales, psicológicas y cognitivas de los alumnos y, en el caso del aprendizaje ubicuo, a todo esto, se suma la necesidad de considerar las condiciones contextuales que pueden influir en el éxito de un aprendizaje colaborativo y situado.

Teniendo en cuenta la problemática planteada, en esta línea de investigación se diseñará un método y se desarrollará el prototipo que genere recomendaciones personalizadas a un estudiante, sobre compañeros con los que puede desarrollar una tarea colaborativamente en un ambiente de aprendizaje ubicuo. El método y el prototipo considerarán para la personalización las condiciones temporales, académicas, disponibilidad horaria, y cercanía física de los estudiantes.

LÍNEA DE INVESTIGACIÓN Y DESARROLLO

Con esta línea de investigación se busca realizar contribuciones teóricas y metodológicas en el campo de la formación de grupos en el aprendizaje colaborativo y ubicuos. En consecuencia, el objetivo general es:

> Analizar, diseñar y construir un prototipo que genere recomendaciones de compañeros con los que un estudiante puede conformar un grupo de trabajo para desarrollar una tarea colaborativamente en un ambiente de aprendizaje ubicuo.

Además, se han definido los siguientes objetivos específicos:

1. Analizar y evaluar los distintos métodos para recomendación de grupos de trabajo colaborativo y ubicuo.
2. Definir los parámetros de personalización que se considerarán para la conformación del grupo.
3. Redefinir el o los métodos para que contemple los parámetros de personalización definidos.
4. Diseñar, construir y evaluar el prototipo que genere grupos de trabajo colaborativo y ubicuo.

Con el propósito de dar cumplimiento a los objetivos planteados, se realizarán las siguientes actividades:

1. *Exploración e investigación bibliográfica sobre técnicas y métodos utilizados actualmente para el modelado y construcción de aplicaciones que recomienden grupos de trabajo colaborativos y ubicuos*: a partir del

estudio exploratorio se desarrollarán los marcos conceptual y metodológico que permitirán abordar el diseño y la construcción del prototipo que genera recomendaciones de compañeros para realizar actividades en forma colaborativas y en ambientes de aprendizaje ubicuo. Se seleccionará la metodología, técnicas y herramientas para el desarrollo del mismo. Se utilizarán, principalmente, las técnicas de estudio de documentación, revisión y fichaje bibliográfico.

2. *Identificación de los requisitos del prototipo y de los parámetros de personalización que se considerarán para la conformación del grupo*: se identificarán y analizarán los requisitos de la aplicación y la posibilidad de aplicar las nuevas técnicas, métodos y estrategias desarrolladas o potencializados, para su construcción. Además, se definirán los parámetros que se considerarán para la conformación de los grupos personalizados. Para el estudio de los requerimientos de estas aplicaciones se usarán técnicas y métodos de la Ingeniería de Requisitos.
3. *Redefinir el o los métodos para que contemple los parámetros de personalización definidos*: a partir del estudio de las estrategias de personalización y de aprendizaje colaborativo, y en base a los parámetros definidos en la actividad anterior se redefine el o los métodos para que contemple todos los parámetros contextuales.
4. *Diseño del prototipo:* se realizará el modelado de requisitos, se construirá el modelo de análisis y el modelo de diseño de la aplicación a construir. Para ello se utilizarán técnicas provenientes de la Ingeniería de Software, de la Ingeniería Web e Inteligencia Artificial.
5. *Construcción del software*: con los modelos de diseño, obtenidos de la actividad anterior, se procederá a desarrollar, en los lenguajes de programación pertinentes, e implementar, en dispositivos de computación ubicua, la aplicación (software) construida. Se realizarán las pruebas de unidad y las pruebas de integración para corroborar la correcta construcción de la aplicación. En base a los resultados de la evaluación se realizarán ajustes en los módulos y/o en la integración de los mismos, si fuera necesario. Para la construcción del software se aplicarán técnicas de programación de dispositivos móviles y de programación web.
6. *Operacionalización de las variables nivel de satisfacción del usuario y desempeño de aplicaciones ubicuas*: a fin de poder evaluar el nivel de satisfacción del usuario y el desempeño de la aplicación ubicua, se identificarán indicadores, presentes en el producto y en el uso del sistema. Para ello, se investigarán estándares referidos a la usabilidad del software, métricas del producto y antecedentes de otros trabajos en los que se evalúen estas variables. Se seleccionarán las métricas más apropiadas y luego, se diseñará el método y las herramientas de recopilación de los datos que permitan medir cada uno de los indicadores definidos.
7. *Ejecución de las aplicaciones en contextos reales*: se implementarán la aplicación construida en contextos reales; y se aplicarán los procedimientos de recopilación de datos definidos en la actividad anterior.
8. *Análisis de resultados*: se procesarán los datos recolectados calculando las métricas previamente definidas. Se elaborará una representación tabular y gráfica de los resultados obtenidos, acompañando los mismos con un análisis interpretativo de los resultados.

4. RESULTADOS ESPERADOS

Los resultados que se esperan obtener de esta investigación, conforme a los objetivos, son:

- Método para formación de grupos que considere los distintos parámetros de personalización, tales como: condiciones temporales, académicas, disponibilidad horaria y cercanía física.

- Prototipo que genere recomendaciones de compañeros con los que un estudiante puede conformar un grupo de trabajo para desarrollar una tarea colaborativamente en un ambiente de aprendizaje ubicuo.

5. FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo de esta línea se integra por un Becario del Programa de Beca EVC-CIN, un Director de la Beca y Director del Proyecto de Investigación indicado en el contexto.

En esta línea de investigación, el becario desarrollará su trabajo final para optar por el título de Licenciado en Sistemas de Información de la Universidad Nacional de Santiago del Estero.

6. BIBLIOGRAFÍA

Amara, S., & Santos, A. (2016). Group Formation in Mobile Computer Supported Collaborative Learning Contexts: A systematic literature review, (April).

Burke, A. (2011). Group Work: How to Use Groups Effectively, 11(2), 87–95.

Cruz, W. M., & Isotani, S. (2014). Group Formation Algorithms in Collaborative Learning Contexts: A Systematic Mapping of the Literature Group Formation Algorithms in Collaborative Learning Contexts: A Systematic Mapping of the Literature, September. Disponible en: https://doi.org/10.1007/978-3-319-10166-8

Dillenbourg, P. (1999). What do you mean by collaborative learning? To cite this version: HAL Id : hal-00190240 What do you mean by ' collaborative learning '?

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design . To cite this version : HAL Id : hal-00190230.

Durán Elena, Unzaga Silvina, Álvarez Margarita, Salazar Nevelin, Gonzalez, Gabriela, Fernández Reuter Beatriz y Zachman Patricia P (2017). "Métodos y Técnicas para desarrollos de Aplicaciones Ubicuas". WICC 2017. Ciudad Autónoma de Buenos Aires. 27 y 28 de Abril. ISBN: ISBN 978-987-42-5143-5. Pag. 838-843. Disponible en: http://wicc.it.itba.edu.ar/pdf/Libro_WICC_2017_ISBN_978-987-42-5143-5.pdf.

Graf S. y Kinshuk (2008). "Adaptivity and Personalization in Ubiquitous Learning Systems". A. Holzinger (Ed.): USAB 2008, LNCS 5298, pp. 331–338, 2008. Springer-Verlag Berlin Heidelberg.

Saadiah Y., Erny A. & Kamarularifin A. J. (2010). "The definition and characteristics of ubiquitous learning: A discussion". International Journal of Education and Development using Information and Communication Technology. (IJEDICT), Vol. 6, Issue 1, pp. 117-127.

Won-Ik Park, Jong-Hyun Park, Young-Kuk Kim and Ji-Hoon Kang. An Efficient Context-Aware Personalization Technique in Ubiquitous Environments. (2010). ICUIMC '10 Proceedings of the 4th International Conference on Uniquitous Information Management and Communication. Article No. 60 ACM. New York, USA. ISBN: 978-1-60558-893-3.

REFORTICCA: Recursos para el Empoderamiento de FORmadores en TIC, Ciencias y Ambiente

María José Abásolo[12], Maria L. Castro[3], Graciela Santos[4], Cecilia Sanz[1], María José Bouciguez[4], Mayra Garcimuño[3], Gabriela Cenich[4], Andrea Miranda[4], Cecilia Papini[4]

[1]Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata (UNLP)
{mjabasolo, csanz}@lidi.info.unlp.edu.ar

[2]Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CICPBA)

[3] Centro de Investigaciones en Física e Ingeniería del Centro de la Provincia de Buenos Aires - CIFICEN (CONICET/CIC/UNICEN)

[4]Educación en Ciencias con Tecnología (ECienTec)
Facultad de Ciencias Exactas - Universidad Nacional del Centro Pcia. Bs. As.
{nsantos, mlcastro, amiranda, gabcen, ,mcpapini}@exa.unicen.edu.ar

Resumen

La línea de investigación y desarrollo que aquí se presenta tiene por objetivo brindar 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" Proyecto de Innovación y Transferencia en Áreas Prioritarias de la Provincia de Buenos Aires de la CICPBA, el cual se lleva a cabo desde 2017 entre los grupos de investigación III-LIDI (UNLP), EcienTec y CIFICEN (UNICEN) de las diferentes áreas involucradas (Abásolo et al, 2017; Abásolo et al, 2018).

Introducción

Por una parte los alumnos, nativos en la era digital, tienen un acercamiento natural a las Tecnologías de Información y Comunicación (TIC). Por otra parte los docentes hacen poco uso de la tecnología en la enseñanza, en particular para la enseñanza de Matemática, Física o Ambiente (Cenich et al, 2017) (Arriassecq 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 hace necesario que los docentes se apropien de la tecnología y adquieran 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 dispositivos para la formación contínua de docentes 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.

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).

Recientemente la RA se popularizó con el uso de aplicaciones móviles que se ejecutan en tablets o celulares, siendo las más populares las que utilizan:

- Códigos QRCode y códigos de barra, para mostrar información relacionada o visitar sitios web
- Imágenes impresas, como los marcadores de RA o imágenes en general, para mostrar información y modelos 3D o videos registrados espacialmente con los marcadores

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)

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 del tipo Google Cardboard[1]. 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 Expeditions[2] que es una aplicación para Android, que permite explorar sitios de interés desde el salón de clase utilizando las gafas Google Cardboard.

Por otro lado, es cada vez más frecuente la inclusión de actividades basadas en RA y RV en el ámbito de las exhibiciones interactivas de ciencia y los museos (Jung et al, 2016). Estas herramientas favorecen el aprendizaje en estos entornos de educación

[1] https://vr.google.com/cardboard/

[2] https://edu.google.com/expeditions/

no formal dado que impactan de manera significativa en la experiencia de los visitantes.

Simulaciones y Juegos

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 son una práctica cultural generalizada entre los jóvenes, lo que los convierte en un medio ideal para el diseño de procesos de aprendizaje. Para diseñar tecnologías educativas que brinden apoyo docente, primero debemos entender las prácticas desarrolladas por los jóvenes con computadoras y especialmente con videojuegos (Bouciguez et al, 2014).

Los videojuegos y simulaciones se emplean en la enseñanza de las ciencias describen algunos videojuegos educativos para enseñar Física (Bouciguez et al, 2013) (Santos, 2016) (Paoletti et al, 2017). Estas herramientas se presentan en el contexto de la enseñanza de las ciencias como herramientas mediadoras con características que benefician la construcción del conocimiento científico. Teniendo en cuenta los propósitos de cada uno de estos entornos interactivos se busca una unificación que combine las características de los mismos a partir de los principios del aprendizaje constructivista en pos de generar entornos interactivos educativos para el aprendizaje de ciencias en contextos educativos (Bouciguez et al. 2019).

Varios autores reconocen que el proceso de desarrollo de un videojuego educativo es complejo, lento y costoso (Westera et al. 2003; Nadolski et al. 2007) ya que implica integrar componentes lúdicos, tecnológicos y educativos, lo cual requiere un trabajo articulado y continuo en equipos interdisciplinarios. Un framework aporta a reducir la complejidad del diseño al proponer una descripción coherente de los componentes básicos y sus interrelaciones, presentado en un nivel general y teórico de razonamiento aplicable en la práctica del diseño de juegos educativos.

Para medir la calidad de un videojuego educativo diseñado se consideran la jugabilidad (González Sánchez, 2010) ya que es importante considerar además de las características funcionales, inherentes a todo sistema interactivo, los atributos y propiedades no funcionales, relativas a la experiencia del jugador al jugarlo.

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).

Comunidad virtual para la formación continua

El dispositivo de formación continua se construye sobre la base del trabajo colaborativo de los equipos docentes con el objetivo de elaborar propuestas de aula que integren tecnologías, denominadas objetos de enseñanza (OE). Los supuestos que le dan sustento al dispositivo son:

-Los docentes necesitan “apropiarse de los modos de hacer” con tecnología y adquirir las competencias necesarias para integrar las tecnologías a la educación en ciencias con usos significativos y pertinentes de las TIC (Rabardel y Bourmaud, 2003; Rodríguez y Barreiro, 2017).

-El trabajo colaborativo para la co-construcción de actividades contribuye a mejorar la "profesionalización docente”, mediante el intercambio de experiencias y debate de ideas con sus pares (Edelstein, 2013).

Un Objeto de Enseñanza (OE) es una propuesta de enseñanza que reúne y compone el conocimiento obtenido en las

experiencias de aula por el equipo docente con el objetivo de ser un insumo para la enseñanza que los profesores podrán adaptar a cada situación áulica. Se lo piensa como recurso dinámico en el marco del modelo TPACK (Koehler; Mishra, 2008), que a partir de una propuesta inicial se enriquece con el análisis de la contribución de recursos TIC al aprendizaje en el caso específico, la elaboración de posibles actividades, estrategias de implementación y la reflexión de la práctica docente.

Para las interacciones entre los grupos se ofreció un espacio virtual en la plataforma Moodle de la Facultad de Ciencias Exactas de la UNICEN, y también se utilizaron grupos y videoconferencias por Whatsapp.

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
- Dispositivos de formación continua de docentes para adquisición de competencias digitales para la integración de tecnologías a la enseñanza.

Resultados y Objetivos

Durante el año 2018 se ha avanzado en los siguientes resultados y objetivos:

- Se diseñó el sitio web del proyecto para la difusión de materiales y recursos educativos que se generen durante el desarrollo.[3]
- Se formó una comunidad virtual de 27 docentes para intercambiar experiencias y abordar en colaboración la producción de OE.
- Se realizaron cinco encuentros de trabajo con los equipo docentes por ciudad; una jornada de trabajo de todos los equipos docentes en febrero de 2019, con el objeto de compartir las producciones y reflexionar sobre las puestas en aulas.
- Se crearon cinco OE, dos se implementaron en aula con el uso de celular y se registró el trabajo de los los grupos de estudiantes y la puesta en común; mientras se prevé la implementación de los tres restantes en el periodo lectivo actual.
- Se diseñó una metodología para generar OE que integren las TIC en las situaciones áulicas; cambio de actitud en los docentes para diseñar propuestas de enseñanza con TIC; profundización del contenido desde las dimensiones disciplinar, tecnológico y pedagógico; mayor compromiso con la tarea y valorización de las prácticas docentes por los resultados obtenidos y el reconocimiento de la comunidad.
- 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 de docentes en diferentes encuentros mediante el llenado del formulario diseñado en Google Forms. 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

[3] https://reforticca.info.unlp.edu.ar/

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 2018), en el Parador Cultural en la Playa, actividad de verano organizada por la Secretaría de Cultura de UNICEN (Quequén, enero 2018) y en la presentación de Estación Ciencias, organizado por el Ministerio de Ciencia y Tecnología de la Provincia de Buenos Aires, en la ciudad de Juárez (junio 2018).
- Se compartieron los lineamientos y las actividades del proyecto en el evento internacional Public Communication of Science and Technology Conference 2018, que tuvo lugar en la ciudad de Dunedin, Nueva Zelanda, entre los días 3 y 8 de Abril de 2018.
- Se diseñó e implementó el sitio web del circuito de reconocimiento de especies nativas conocido como Sendero Pampa.[4] Se incorporó RA en formato de códigos QR en la señalización de 10 estaciones del mencionado circuito.
- Se diseñó un vídeo referido a los polinizadores locales con la técnica *stop-motion* para incorporar al Sendero Pampa. Actualmente en etapa de edición.
- Se desarrolló el Proyecto RA Gliptodonte utilizando *Aumentaty Creator* y *Scope Aumentaty* que acompañó la exhibición de fósil de gliptodonte neosclerocalyptus hallada en el partido de Tandil y que se expone en la sala de paleontología en el Museo Tradicionalista del Fuerte Independencia Tandil. Se presentó durante vacaciones de invierno 2018; y posteriormente quedó para su uso en la sala mencionada.
- Se dictó el taller "RA y RV: sus posibles usos en el aula" para docentes de escuelas primarias y secundarias (septiembre 2018). Asistieron más de 60 docentes. Se propone actualmente continuar la formación e intercambio en un espacio virtual de enseñanza y aprendizaje.
- Se desarrolló el sitio "KitLab. LAboratorio de Ciencias"[5] para acompañar los carros de laboratorio entregados a las escuelas primarias por el Plan Nacional Ciencias Naturales del Ministerio de Educación. El sitio tiene por objeto dar a conocer a los docentes los elementos que componen un kit de química distribuido. Mediante la captación de un código QR pegado a cada elemento del kit se puede acceder a la información relacionada en el sitio. Como complemento se diseñó un póster con las imágenes de los elementos del kit de laboratorio, los cuales actúan como marcadores de RA y se aumentan con vídeo y links al sitio web..

Formación de recursos humanos

El desarrollo de algunos recursos educativos fue realizado por alumnos dentro del marco de la carrera Profesorado de Informática de la Facultad de Cs. Exactas de la UNICEN y alumnos del posgrado de la Facultad de Informática de la UNLP.

[4] ttps://www.senderopampa.com.ar/

[5] http://kitlab.exa.unicen.edu.ar/

En la actualidad hay en curso diferentes tesis de postgrado, grado y becas de entrenamiento en el marco de esta línea de investigación:

- 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 Ciencias, Fac. Inf. de la UNLP, 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 de Ing. de Sistemas, Fac. Cs. Exactas de la UNICEN, en curso).

Referencias

Abásolo, M.J.; Sanz, C.; Santos, G.; Castro, M.; Miranda, A.; Cenich, G.; Bouciguez, M.J. ; Papini, C. (2018) *REFORTICCA: Recursos para el Empoderamiento de FORmadores en TIC, Ciencias y Ambiente.* XX Workshop de Investigadores en Ciencias de la Computación (WICC 2018), pp. 1197-1202, RedUNCI, ISBN 978-987-3619-27-4

Abásolo, M.J.; Sanz, C.; Naiouf, M.; De Giusti, A.; Santos, G.; Castro, M.;Bouciguez, M.J. G. (2017) *Realidad Aumentada, Realidad Virtual e Interacción Tangible para la Educación.* XIX Workshop de Investigadores en Ciencias de la Computación (WICC 2017), pp. 1312-1316, RedUNCI, ISBN 978-987-42-5143-5

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

Arriaseqc, I.; Santos, G. (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 - .ISNN 2346-8866

Artola,Verónica; Sanz,Cecilia; Moralejo, Lucrecia; Pesado, Patricia; Baldasarri,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

Bouciguez, M.J., Braunmüller, M., Bravo, B., Santos, G.y Abásolo, M. J. (2019) *Desarrollo del videojuego "SpaceEscape: The F.E.M."para una secuencia didáctica de inducción electromagnética*. X Congreso Iberoamericano de Educación Científica Enseñanza y Aprendizaje de las Ciencias en Debate. Montevideo, Uruguay, 25, 26-28 de Marzo del 2019

Bouciguez, M. J.; Santos, G. ;Abásolo, M. J. (2014) *Towards the use of video games for learning: a survey about video games preferences of Engineering.* Journal of Computer Science & Technology; vol. 14, no. 1, p. 25-31, ISSN: 1666-6038

Bouciguez, M. J.; Santos, G. ;Abásolo, M. J. (2013) *Potencialidad de los videojuegos en el aprendizaje de Física.* Actas de WEFA 2013 I Workshop de Enseñanza de Fisica en Argentina, Universidad Nacional del Centro de la Pcia. de Bs. As., ISBN 978-950-658-342-2.

Cenich, G.; Araujo, S.; Santos, G. (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

Edelstein, G. (2013) *Formar y formarse en la enseñanza*. Buenos Aires: Editorial Paidos.

Gavilanes, W., Abásolo, M. J., Cuji, B. (2018) *"Realidad Aumentada en la Educación: una Revisión desde la Perspectiva Pedagógica",* Revista ESPACIOS. ISSN 0798 1015

González Sánchez, J. L. (2010). Tesis doctoral. Jugabilidad. Caracterización de la experiencia del jugador en videojuegos. Granada: Universidad de Granada.Hernández

Ortega, J.; Pennesi Fruscio, M.; Sobrino López, D. y Vázquez Gutiérrez, A. (2012) *Tendencias Emergentes en Educación con TIC.* Asociación Espiral, Educación y Tecnología, ISBN: 978-84-616-0448-7

Koehler, M.; Mishra, P. (2008), "Introducing TPCK", en AACTE Committee on Innovation and Technology (Eds.), Handbook of technological pedagogical content knowledge (TPCK) for educators, New York, Routledge, pp. 3-30.

Jung T., Tom Dieck M.C., Lee H., Chung N. (2016) *Effects of Virtual Reality and Augmented Reality on Visitor Experiences in Museum.* En: Inversini A., Schegg R. (eds) Information and Communication Technologies in Tourism 2016. Springer, Cham

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

Nadolski R., Hummel H., van den Brink H., Hoefakker R., Slootmaker A., Kurvers H. & Storm J. (2007) *Emergo: methodology and toolkit for efficient development of serious games in higher education.* Available at: http://dspace.ou.nl/handle/1820/1046

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

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.

Rabardel, P.; Bourmauda, G. (2003). *From computer to instrument system: a developmental perspective.* Interacting with Computers 15, pp. 665–691.

Rodríguez, M.; Barreiro, P. (2017) Consideraciones sobre la formación de profesores de matemática y su apropiación de las nuevas tecnologías. En Cabello y López (eds.) Contribuciones al estudio de procesos de apropiación de tecnologías. Rada Tilly: Del Gato Gris; Ciudad Autónoma de Buenos Aires: Red de Investigadores sobre Apropiación de Tecnologías

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/120835177/issue

Sanz, C. & Artola, V. (2017) *Folleto de Autorregulación.* Vídeo disponible en: 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

Westera W., Hommes M.A., Houtmans M. & Kurvers H.J.(2003) *Computer-supported training of psychodiagnostic skills.* Interactive Learning Environments 11, 215–231.

UAI CASE: DESARROLLO Y EVALUACIÓN DE COMPETENCIAS EN LA INGENIERÍA DE SOFTWARE EN UN ENTORNO VIRTUAL DE APRENDIZAJE COLABORATIVO

Nicolás Battaglia, Carlos Neil, Marcelo De Vincenzi, Juan Pablo Beltramino

CAETI - Centro de Altos Estudios en Tecnología Informática
Facultad de Tecnología Informática. UAI - Universidad Abierta Interamericana
{nicolas.battaglia, carlos.neil, medevincenzi}@uai.edu.ar,
juanpablo.beltramino@alumnos.uai.edu.ar

RESUMEN

Los avances en Tecnología de Información y Comunicación generan grandes aportes a la educación facilitando, entre otras cosas, que los alumnos puedan acceder a la información de forma ubicua. Esta característica permite que las entidades educativas puedan desarrollar ofertas educativas, con un alto componente de virtualidad a partir de nuevas disposiciones ministeriales, en modalidad presencial

Si bien el desafío principal consiste en diseñar una infraestructura para poder brindar servicios de educación mediada por la tecnología es necesario, también, permitir que los alumnos puedan trabajar de forma colaborativa y optimizar al máximo el proceso formativo propuesto por esta modalidad.

La entornos de educación a distancia permiten también cambios en los procesos de Enseñanza y Aprendizaje, entre ellas el Aprendizaje Basado en Competencias y centrado en el estudiante, dando lugar a nuevos desafíos: diseñar entornos virtuales colaborativos para aprendizaje basado en competencias y, también, modificar y adaptar los criterios y estrategias de evaluación motivados por esta disrupción

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 que brinde un espacio virtual de colaboración para el proceso formativo basado en competencias.

Palabras clave: Aprendizaje Colaborativo, Educación por Competencias, Ingeniería de Software, Trabajo Colaborativo, UML.

CONTEXTO

Este proyecto de investigación se desarrolla 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 asignaturas 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 con un nivel de complejidad que se va incrementando en los sucesivos años conforme incrementan su nivel de competencia.

Cabe destacar que este proyecto, se contextualiza en el documento recientemente publicado por el Consejo Federal de Decanos de Ingeniería (CONFEDI) llamado "Propuesta de Estándares de Segunda Generación para la Acreditación de Carreras de Ingeniería en la República Argentina" [1], en el cual se incorpora el concepto de competencia, describiendo dentro de las condiciones

curriculares comunes las Competencias de Egreso que un alumno debe acreditar para acceder a su graduación.

1. INTRODUCCIÓN

La innovación constante en las Tecnologías de Información y Comunicación (TIC) ha modificado notablemente diferentes aspectos sociales y culturales, y la educación es uno de los ámbitos que más impacto y aportes recibió de esta disrupción tecnológica. La posibilidad de contar con dispositivos que permiten acceder a la información en cualquier momento del día, que poseen herramientas tales como *chats*, foros, *blogs* y el acceso a redes que promueven la interacción entre personas, posibilita un cambio de enfoque en las metodologías de enseñanza y aprendizaje [2] .

Tal es así que el trabajo colaborativo, las TIC y la enseñanza se integran para crear entornos de aprendizaje colaborativo asistido por computadora (CSCL). Esta integración se basa en el impacto social del trabajo colaborativo y las TIC en la educación. Este impacto transforma la manera tradicional de enseñanza, con varios cambios, entre ellos el Aprendizaje Basado en Competencias y centrada en el alumno y el impacto de las TIC en los procesos de enseñanza y aprendizaje para lo cual es necesario también modificar y adaptar los criterios y estrategias de evaluación motivados por esta disrupción.

Por un lado, el trabajo colaborativo apoyados en el concepto de Zona de Desarrollo Próximo (ZDP) propuesto por Vygotski a principios del siglo XX [3] y, por el otro, las TIC en el marco de las propuestas de *e-learning* o *blended learning*. El desarrollo creciente de las TIC junto con el concepto de trabajo colaborativo, conforman los entornos de Trabajo Colaborativo Asistido por Computadora (*Computer Supported Collaborative Work*, CSCW). Este concepto, integrado en entornos de enseñanza y aprendizaje colaborativo, da origen a los entornos de Aprendizaje Colaborativos Asistidos por Computadora (*Computer-Supported Collaborative Learning,* CSCL).

Los nuevos estándares para la Acreditación de Carreras de Ingeniería en la República Argentina (Libro Rojo de CONFEDI) [1], dan cuenta de la necesidad de que los entornos CSCL tengan en consideración el Aprendizaje Basado en Competencias y centrado en el estudiante, además de todas las herramientas de seguimiento y evaluación que sean requeridas por este paradigma (figura 1)

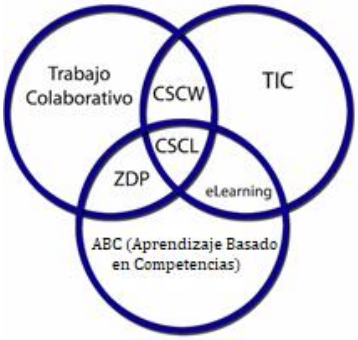


Figura 1. Conformación del modelo CSCL

El estudio teórico de la Ingeniería de Software (IS) no es suficiente para comprender y resolver los problemas de cooperación y colaboración que surgen durante el desarrollo de un proyecto informático. Los estudiantes suelen centrar su esfuerzo en aspectos técnicos y asumen que los inconvenientes vinculados al trabajo en equipo no impactarán en el proyecto [4]. En virtud de resolver el problema planteado previamente, Daniele et al. [5] proponen, para mejorar los aspectos comunicacionales, la integración de las plataformas CSCL con entornos especializados en la resolución de problemas prácticos y técnicos en un proyecto de desarrollo de software; además, esta propuesta potencia el aprendizaje de la IS y permite el desarrollo de competencias profesionales y sistemáticas en el desarrollo de software [6].

De la misma manera que sucede en los procesos presenciales o tradicionales de enseñanza, en los entornos CSCL la evaluación cumple un papel muy importante. Según el estudio realizado en un relevamiento previo [7], en el modelo de enseñanza y aprendizaje de la IS mediado por tecnología, los estudiantes cumplen un papel de mayor protagonismo y responsabilidad y, en consecuencia, los docentes asumen un nuevo

rol cuando es el alumno quien adquiere mayor autonomía en su proceso de aprendizaje, a través de entornos virtuales, autonomía que es una característica distintiva de la enseñanza centrada en el estudiante

En este contexto, la evaluación es entendida como un proceso que promueve el aprendizaje con una finalidad formativa, más que como un proceso de control de resultados. Por consiguiente, un proceso específico de enseñanza y aprendizaje en un entorno virtual colaborativo requiere también de un proceso de evaluación, autoevaluación y seguimiento acorde al modelo de aprendizaje colaborativo planteado por el concepto CSCL [8]. El énfasis en la educación centrada en el estudiante permite retomar el modelo de aprendizaje basado en competencias, el cual también promueve reformular y adaptar los criterios y estrategias de evaluación motivados por disrupción tecnológica que da origen a los entornos CSCL.

En este proyecto nos enfocaremos en la enseñanza y el aprendizaje del modelado de software y utilizaremos el estándar UML como lenguaje para especificar, construir, visualizar y documentar, para lo cual propondremos el rediseño de un entorno CSCL específico para dicha área de conocimiento, con un enfoque centrado en el Aprendizaje Basado en Competencias

2. uCASE-CL

Los modelos de aprendizaje colaborativo asistido por computadora están inmersos en un universo tecnológico ubicuo, basado en las tecnologías de la comunicación. Este tipo de modelo fue denominado uCSCL por Coto, Collazos y Rivera [9], entre otros autores [6]. El modelo CSCL está compuesto por tres dimensiones (Ciencias de la Computación, Psicología y Pedagogía), a estas se agrega una cuarta dimensión que representa la tecnología subyacente que materializa el concepto de ubicuidad [10].

La propuesta de uCASE-CL [8] [11] [12] fue ampliar el modelo CSCL mencionado para contar con un bloque funcional de evaluación integrando satisfactoriamente las herramientas de enseñanza y aprendizaje necesarias para áreas de conocimiento específico, como sucede con la Ingeniería de Software.

El modelo uCASE-CL propuesto no considera las herramientas necesarias para poder adoptar el sistema aprendizaje basado en competencias. Si bien está diseñado para centrar la educación en el alumno, no permite definir competencias ni tampoco brinda herramientas de evaluación y seguimiento, para poder obtener indicadores particulares sobre la evolución del aprendizaje y las competencias vinculadas.

Por lo tanto, el objetivo principal de este proyecto es estudiar e investigar diferentes técnicas que permitan integrar la educación basada en competencias dentro de entornos virtuales colaborativos, a fin de aprovechar las bondades de las TIC y el concepto de ubicuidad para brindar información en tiempo real sobre la evolución del aprendizaje en cada una de las competencias que estén vinculadas con el objeto de estudio.

3. APRENDIZAJE BASADO EN COMPETENCIAS

La universidad debe preparar al estudiante en un entorno de competencias genéricas y específicas que les permitan seleccionar información y afrontar los problemas que plantea esta sociedad cambiante.

McClelland [13], en la búsqueda de una alternativa a las pruebas de aptitud e inteligencia tradicionales, desarrolla el concepto de “competencia” definido como una característica subyacente de una persona que le permite demostrar un desempeño superior en un determinado puesto, rol o situación, haciendo la diferencia entre personas con desempeño excelente versus personas con desempeño promedio [14]. Por otro lado, el concepto de competencia engloba todo un conjunto de conocimientos, procedimientos y actitudes que se combinan, coordinan e integran, para que el individuo logre «saber», «saber hacer», «saber ser» y «saber estar». El dominio de estos saberes le permiten ser «capaz de» actuar con eficacia y eficiencia en situaciones que se le presenten [15]. En este

sentido, la competencia no reside en los recursos (capacidades) sino en la movilización misma de los recursos. Para ser competente es necesario poner en juego el repertorio de recursos [15]. Además, el aprendizaje por competencias ha ocasionado toda una renovación de las teorías psicopedagógicas de los aprendizajes, dando lugar a la evolución de los esquemas de referencia de la formación de profesionales [15].

Es importante destacar que la capacitación y educación basada en competencia, es una metodología de instrucción que identifica las habilidades básicas, conocimientos y actitudes que satisfacen normas específicas, enfatiza estándares de ejecución y facilita el aprendizaje individual [14]. Por último, la formación de individuos bajo un modelo de competencias, les permite a estos estar preparados para responder de manera autónoma e integral, y también para adaptarse a procesos que se actualizan permanentemente [15].

4. EVALUACIÓN Y COMPETENCIAS

La evaluación al igual que cualquier otra actividad de estas características requiere de interacción entre las personas involucradas, incluyendo docentes y alumnos participantes de un grupo de trabajo [6].

Por otro lado, el concepto de evaluación para el aprendizaje colaborativo puede definirse como un proceso global, continuo, contextualizado, planificado, interactivo, y estratégico que permite identificar, comprender, valorar y reorientar tanto la evolución del aprendizaje y sus potencialidades, como la reflexión sobre las experiencias de aprendizaje, compartidas en una situación didáctica de grupo colaborativo [10].

El ABC requiere un sistema de evaluación variado, debido a que cada competencia tiene componentes muy distintos que necesitan procedimientos diversos para ser evaluados correctamente [19]. La evaluación basada en competencias se basa en el acceso a fuentes múltiples y variadas de información con el fin de determinar si los estudiantes han alcanzado el nivel esperado de desarrollo de competencias, así como un grado suficiente de dominio de los recursos vinculados a cada competencia [16].

Si bien existen diversas metodologías de evaluación de competencias [17], en la actualidad uno de los mejores instrumentos, son las rúbricas, debido a su facilidad de uso y a la oportunidad de obtener aspectos complejos, imprecisos y subjetivos [7] [18].

5. 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.

6. RESULTADOS OBTENIDOS/ESPERADOS

6.1 Resultados obtenidos:

- Desarrollo, diseño e implementación de una herramienta CASE colaborativa ubicua multiplataforma, denominada UAI Case, integrada con los conceptos de aprendizaje basado en competencias.
- 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.

6.2 Objetivos Futuros / Resultados esperados:

- Diseñar herramientas que permitan integrar el concepto de aprendizaje basado en competencias junto con herramienta que permitan evaluar y realizar seguimientos en entornos colaborativos como es UAI Case.
- Generar indicadores en tiempo real que permitan medir la evolución del conocimiento

durante el aprendizaje basado en competencias.

• 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.

• Desarrollar sistemas de tutorías inteligentes mediante la implementación de tutores cognitivos que imiten el rol del profesor, guíen el desarrollo del aprendizaje y ofrezcan pistas a los estudiantes cuando están atascados en un problema

• Utilizar analíticas del aprendizaje que mediante la medición, recopilación y análisis de datos de los estudiantes durante el proceso de enseñanza aprendizaje, permita detectar los errores más comunes y proporcionen una respuesta en tiempo real al estudiante.

7. 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 6 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.

8. BIBLIOGRAFÍA

1. CONFEDI (2018): Propuesta de estándares de segunda generación para la acreditación de carreras de ingeniería en la República Argentina “Libro Rojo de CONFEDI”.
2. Lavigne, G., Vasconcelos Ovando, M. P., Organista Sandoval, J., & McAnally Salas, L. (2012). Exploración preliminar del aprendizaje colaborativo dentro un entorno virtual. Revista Electrónica" Actualidades Investigativas en Educación", 12(3).
3. Vygotsky, L. S. (1997). The collected works of LS Vygotsky: Problems of the theory and history of psychology (Vol. 3). Springer Science & Business Media
4. Bouillon, P., Krinke, J., & Lukosch, S. (2005, April). Software engineering projects in distant teaching. In 18th Conference on Software Engineering Education & Training (CSEET'05) (pp. 147-154). IEEE.
5. 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.
6. Battaglia, N., Neil C., (2017) *Integración de una Herramienta CASE en un Entorno Académico Colaborativo para la Enseñanza de Ingeniería de Software* (Tesis de Maestría). Universidad Abierta Interamericana. Buenos Aires. Argentina.
7. Battaglia, N. Martínez, R. Otero, M., Neil, C., De Vincenzi M. (2016). Autoevaluación Colaborativa por medio de Rubricas en Entornos Virtuales de Enseñanza y Aprendizaje. In I Workshop sobre Innovación en Centros Educativos y de Investigación (I WICEI), II Jornadas Argentinas de Tecnología, Creatividad e Innovación (JATIC)
8. Battaglia, N., Neil, C., De Vincenzi, M., Martínez, R., González Dana. (2017, junio). uCASE-CL: Aprendizaje Colaborativo de la Ingeniería de Software en Entornos Virtuales Ubicuos. In XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2017).
9. 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)
10. Neil Carlos, Battaglia N., Beltramino J., Zalazar A. Evaluación de modelos UML en Entornos Virtuales Colaborativos. CIITI 2018.
11. 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).
12. Battaglia, N., Neil C., Martínez R., De Vincenzi M. (2018). UAI Case: An uCASE-CL Model Implementation. In the 2018 International Conference On Creative and Innovative Technology in Education (i-CITE). In Senai, Johor, Malaysia.
13. McClelland, D. (1973). Testing for competence rather than for 'intelligence'. American Psychologist. 28, 1-14.
14. Leyva, M. R. V. (2008). Diseño curricular por competencias. Asociación Nacional de Facultades y Escuelas de Ingeniería.
15. Fernández, J. T., & Bueno, C. R. (2015). Evaluación de competencias profesionales en educación superior: retos e implicaciones. *Educación XX1*, *19*(1).
16. March, A. F. (2010). La evaluación orientada al aprendizaje en un modelo de formación por competencias en la educación universitaria. *REDU: Revista de Docencia Universitaria*, *8*(1), 11-34.
17. Cubero-Ibáñez, J., Ibarra-Sáiz, M. S., & Rodríguez-Gómez, G. (2018). Propuesta metodológica de evaluación para evaluar competencias a través de tareas complejas en entornos virtuales de aprendizaje. Asociación Interuniversitaria de Investigación Pedagógica (Aidipe) Miembro de la European Educational Research *(EERA)*, *36*(1), 159.
18. Schiter T, Battaglia N., Neil C., (2018) El Porqué de las Rúbricas para la Evaluación y la Autoevaluación. CONAIISI. Mar De Plata, Bueno Aires, Argentina.
19. Villa, A. (2007). Aprendizaje basado en competencias: una propuesta para la evaluación de las competencias genéricas (No. 371.392 V7113a Ej. 1). Ediciones Mensajero.

Un enfoque integral para propiciar cursos abiertos on line desde la UNIVERSIDAD NACIONAL DE CÓRDOBA

Laura Cecilia Díaz Dávila, Silvia Edith Arias, José Daniel Britos, Matías Ramón Cuenca del Rey, Guillermo Edurado Zaballo, José Luis Galoppo, Adolfo Vignoli, Gisela Hirschfeld, Marisa Dasso, Yamil Giralda, María Sofía Sappia, José Manuel Piro, Loreley Nahir Bustamante, Sandro Comerci, María Lourdes Dorado, Elisabeth Leonhardt, Gonzalo Leonel Vilches, Amanda Quenard y Fuentecilla
Departamento de Computación
Facultad de Ciencias Exactas, Físicas y Naturales
Universidad Nacional de Córdoba
Av. Vélez Sarsfield 1611, Córdoba
0351-4334409
laura.diaz@unc.edu.ar, sil_var@hotmail.com, dbritos@unc.edu.ar, mcuenca@unc.edu.ar, guillermo.zaballo@unc.edu.ar, jose.galoppo@unc.edu.ar, adolfo.vignoli@unc.edu.ar, gise320@gmail.com, dasso.marisa@gmail.com, yamilgiralda@gmail.com, sofisappia@gmail.com, josemanuelpiro@alumnos.unc.edu.ar, comercisandro@gmail.com, m.lourdes.dorado@alumnos.unc.edu.ar, eleonhardt@alumnos.unc.edu.ar, gonzalo.vilches@alumnos.unc.edu.ar, amandaquenard@gmail.com

RESUMEN

Este proyecto de investigación aplicada da continuidad a las acciones iniciadas en el bienio anterior. Se orienta a la construcción de un prototipo MOOC, en un contexto ya institucionalizado para su desarrollo. Desarrolla acciones de formación de investigadores en tecnologías novedosas como: procesos de digitalización, blockchain, sistemas software, plataformas educativas y herramientas del campo de la Inteligencia Artificial; para coadyuvar a los diferentes enfoques requeridos en el contexto tecnológico actual caracterizado por sus vertiginosos cambios y el aprendizaje colaborativo. Avanza hacia la construcción de propuestas creativas para atender a la articulación entre niveles medio y universitario, por una parte al Ciclo General de Conocimientos Básicos (CGCB) por otra, específicamente en Matemática para Ingeniería en la Universidad Nacional de Córdoba (UNC). Profundiza en el conocimiento del modelo del estudiante, en sus características socioeconómicas, académicas y cognitivas, mediante el uso de Tecnologías Inteligentes de Explotación de la Información (TIEI), no sólo para mejorar los procesos de aprendizaje y evaluación, sino descubrir patrones de comportamiento relevantes para procesos decisionales en la gestión de Educación Superior (ES). Por último, desarrolla actividades interdisciplinarias que posibilitan mejores servicios a la sociedad con la finalidad última de facilitar la apropiación del conocimiento y la tecnología en ES.

***Palabras clave*: Cursos on line abiertos, Tecnologías Inteligentes, Enfoque multidisciplinar**

CONTEXTO

En la categoría de "Proyecto Consolidar III", esta investigación aplicada para favorecer acceso, continuidad y permanencia, mitigando el desgranamiento; convoca a un equipo consolidado, cuyos antecedentes con diferentes matices datan en más de una década. Además, trabaja articuladamente con dos "Proyectos Formar" de la Facultad de Ciencias Exactas, Físicas y Naturales y un proyecto Consolidar de la Facultad de Ciencias de la Comunicación, en

el marco del Programa "Apropiación Social del Conocimiento y la Tecnología". Todos, acreditados y financiados por la Secretaría de Ciencia y Técnica de la UNC.

1. INTRODUCCIÓN

El objeto de estudio es la ES en contextos de masividad, escenario áulico en el cual la comunicación personal entre el docente y el estudiante se ve impedida en razón del número de participantes.

La finalidad es favorecer la apropiación del conocimiento, acordando con la declaración del Consejo Interuniversitario Nacional (CIN) de 2005, al que se sumó el Consorcio del Sistema de Información Universitaria (SIU) en 2013.

En Argentina, el acceso a las Universidades públicas y el desgranamiento [1] asociado a la masividad despierta el interés por ofrecer propuestas creativas acordes a esta realidad. Alineado con las conclusiones de la Conferencia Regional en Educación Superior de América Latina y el Caribe (CRES) 2008, "La Educación Superior es un bien público social, un derecho humano universal y un deber del Estado. Ésta es la convicción y la base para el papel estratégico que debe jugar en los procesos de desarrollo sustentable de los países de la región…"; en tiempos en que la construcción de saberes constituye un desafío importante [2].

Los MOOC por su sigla en inglés (Massive Open Online Courses), se presentan como estrategia generalizada que emerge como una Universidad paralela en la nube, siendo ejemplos de consorcios que lideran Coursera y edX. Con su advenimiento, los ciudadanos, desde diversos lugares, pueden acceder a niveles de acreditación de Universidades reconocidas [3].

En la región, las redes como Red Clara a nivel Latinoamericano o Innova- Red a nivel nacional, constituyen una evidencia de grandes esfuerzos. Los avances en infraestructura en Latinoamérica y en el fortalecimiento de redes de Internet2 a NREN (National Research and Education Network) [4] son alentadores.

Las características propias de Latinoamérica devienen en una importante responsabilidad social hacia los estudiantes, como lo señala Franqois Vallaeys [5]. Los MOOC de las grandes corporaciones Universitarias vienen avanzando a pasos agigantados [6]. En este sentido Brasil ha dado un paso adelante con Veduca [7], con un gran impacto en la educación de habla portuguesa. En lengua Española apenas se asoma "mooc.es" [8]. Coursera actualmente cuenta con aproximadamente 22 millones de usuarios [9]. En el consorcio OCW se desarrolló la plataforma educativa de uso libre llamada OpenedX [10], fundada por Harvard University y MIT en 2012, es un destino de aprendizaje en línea que ofrece MOOC de alta calidad de las mejores universidades e instituciones a estudiantes de todo el mundo [10] y sin fines de lucro, compuesta por el xConsortium [11].

En este escenario, nuestro equipo aborda un tratamiento integral, aunque no completo, atendiendo a las políticas de la gestión de gobierno de la UNC, los paradigmas de Construcción Colaborativa de Aprendizajes y las Nuevas Estrategias de Evaluación y Acreditación de Saberes.

En primer lugar, la investigación se orienta a Obtener conocimiento del estudiante, principal actor del complejo escenario, a partir del uso de Tecnologías Inteligentes de Explotación de la Información (TIEI), para contribuir: en el diseño de MOOC posibilitando que se adecuen a sus potenciales usuarios, con información útil para los actores directamente involucrados en procesos de aprendizaje, evaluación y en las decisiones institucionales. Los hallazgos en relación al descubrimiento de patrones de aprendizaje bajo la plataforma Moodle, la caracterización socioeconómica en base a la información de SIU Guaraní y la investigación exploratoria del módulo Analytic de OpenedX, alientan a esta línea de Inteligencia Artificial.

Son relevantes los estudios realizados en universidades del país para la identificación de las características socioeconómicas y demográficas más significativas en la construcción de su modelo de comportamiento académico [16] [17] [18]. En síntesis, estas experiencias motivan a profundizar el uso de herramientas de Inteligencia Artificial como aspecto necesario para transitar el camino hacia el diseño de MOOC.

En segundo lugar, se focaliza en la mejora del aprendizaje en Matemática para las carreras de Ingeniería en los procesos de articulación entre los niveles medio y universitario como área temática prioritaria escogida para la incorporación de nuevas tecnologías educativas en cursos masivos, justificado en las experticias de los equipos de investigación en colaboración con este, las demandas detectadas y la consideración de los CGCB y los Consorcio InterUniversidades. En esta línea de investigación en el área crítica de la matemática aportan dos proyectos "Formar" vinculados.

En Argentina, Universidades como la de Buenos Aires, Córdoba, Rosario y la Plata, ofrecen asignaturas de varios miles de estudiantes. Las plataformas tradicionales como Manhattan, Claroline, Moodle y Sakai [22] [23], se han ido quedando atrás en sus servicios, en cuanto a la masividad de la enseñanza y la interactividad que con los videos se refiere.

El principal problema del despliegue de estos cursos voluminosos, en cantidad de alumnos, es la inversión en tiempo y equipamiento; razón por la cual sus implementaciones las lideran consorcios de Universidades. Así, el "Consorcio UBA-UNC en la Sociedad del Conocimiento" desde 2017, apunta a articular esfuerzos para complementarse en el campo académico, científico y extensionista.

En tercer lugar, se propone diseñar un prototipo de MOOC en una temática que posibilite visibilidad internacional para la UNC. En este sentido, los cursos de buceo, de corta duración y multidisciplinares ya que involucran conocimiento de física, fisiología y medicina; admiten certificación centralizada por organismos internacionales como la Federación Argentina de Actividades Subacuáticas (FAAS) resultan adecuados. Se espera receptividad internacional si los incorpora la Confederación Mundial de Actividades SubAcuáticas (CMAS). Existen antecedentes de entidades como ScubaSchools International (SSI), Asociación profesional de instructores de buceo Professional Association (PADI) y ScubaDiving International (SDI), que han desarrollado cursos online de tal forma que la instrucción que reciban sus alumnos en las distintas escuelas alrededor del mundo posea el mismo nivel académico.

La CMAS posee escuelas adheridas en todo el mundo, pero todavía no ha incursionado en la educación online. Frente a esta oportunidad estratégica para desarrollar MOOC, se tomó contacto con la FAAS en 2017 (posee 200 escuelas de buceo en territorio Argentino) y se firmó un convenio de cooperación con la UNC.

El desarrollo de experticia para el diseño de cursos abiertos, basada en los procesos de producción, implementación y evaluación de guiones audiovisuales, que favorezcan la accesibilidad y la calidad educativa, **es una sub línea de investigación**. Nuestro equipo de investigación escogió a Open edX como plataforma para gestionar los MOOC, tras una investigación exploratoria que incluyó estudios comparativos, análisis de antecedentes y exploración de sus potencialidades durante el bienio anterior. Alineada con las políticas institucionales actuales de la UNC y la presidencia del CIN desde Marzo de 2018. En octubre de 2017 el consorcio edX le dio la bienvenida a la UNC.

En cuarto lugar, basado en la evidencia que las plataformas que brindan servicios a través de Internet requieren conexiones seguras en sus transacciones. Específicamente, las de aprendizaje y gestión de procesos institucionales, como Moodle, Educativa,

sistemas SIU y Open edX. Exploramos soluciones novedosas: La tecnología disruptiva de blockchain es el objeto de estudio. Muy útil ámbitos institucionales, su aplicación incentivaría entornos educativos más colaborativos e interconectados. Según [14] "el punto de partida es la acreditación del currículum vitae, si bien existen otras aplicaciones a portafolios, evidencias de aprendizajes...". Emerge una tendencia de reconocidas instituciones para trabajar en Blockchain y educación [15]. Compañías como Sony Global Education, Inc. que anunció la incorporación de blockchain, desarrollando tecnología para el intercambio académico y sus registros, abierto y seguro (Tokio, Japón). Es relevante investigar el rol de las Universidades como pares, que requieren capacidades como actores en una base de datos distribuida confiable para compartir de forma abierta el dominio académico y los registros de progreso y otras aplicaciones que emergen como respuesta a las inquietudes de los investigadores en formación de este equipo.

En pocas palabras, la formación en espacios virtuales abiertos, materializada en acciones orientadas a propiciar los MOOC desde la UNC, concebidos como una alternativa de mejora de accesibilidad a Universidades [21].

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

1.Investigación en aspectos relativos a la plataforma Open edX.

2.Producción de un prototipo de MOOC en Open edX con temática en Buceo.

3.Descubrimiento de patrones de comportamiento del estudiante de ES, aplicando Inteligencia Artificial.

4.Investigación aplicada en blockchain: exploración de plataformas, evaluación para diversos ámbitos institucionales.

5.Incorporación de herramientas tecnológicas para matemática: videos cortos, animaciones con html5, innovaciones de entornos virtuales.

6.Formación de grado y posgrado: Educación en entornos virtuales, Blockchain aplicado a Instituciones Educativas, Blockchain aplicado a voto electrónico, Digitalización de expedientes, Inteligencia Artificial para descubrir patrones de comportamiento de estudiantes, Diseño de productos en Open edX y Herramientas, como html5, para animaciones en ES.

3. RESULTADOS OBTENIDOS/ESPERADOS

Si bien este proyecto se enfoca a contribuir con el desarrollo de MOOC desde la UNC para atender el problema de la ES en contextos de masividad; en razón del carácter integral de su abordaje, su impacto merece ser distinguido:

Favorece la visibilidad institucional, en razón del curso de buceo, prototipo MOOC.

Promueve la formación de investigadores en tecnologías novedosas como blockchain, los sistemas software, las plataformas educativas y las herramientas de la Inteligencia Artificial.

Contribuye a paliar las dificultades de la articulación entre el nivel medio y la universidad en Matemática para Ingeniería.

Ofrece un mejor conocimiento del estudiante, de relevancia no sólo para los MOOC, sino también para procesos decisionales en ES.

Por último, el efecto multiplicador que emerge de sumar integrantes de otras especialidades.

Los Objetivos específicos y sus avances:

1. Dar continuidad a las acciones iniciadas en el bienio inmediato anterior: evaluar los recursos audiovisuales generados y dejarlos a disposición de la comunidad. Estado de Avance: En desarrollo, conforme al cronograma del Proyecto.
2. Profundizar en el conocimiento del modelo del estudiante, en sus características socioeconómicas, académicas y cognitivas, mediante el uso de TIEI, atendiendo a los requerimientos que emergen del contexto actual. Estado de avance: Idem anterior
3. Desarrollar un curso de Buceo abierto on line sobre la plataforma Open edX, en

consonancia con la Asociación de la UNC a edX, abarcando diversos enfoques en la construcción de MOOC. Estado de avance: Idem anterior

4. Lograr avances de propuestas innovadoras que atiendan la articulación nivel medio y universitario y el CGCB, en Matemática, para Ingeniería de la UNC. Estado de avance: Idem anterior
5. Favorecer la formación de los Investigadores en Formación del equipo en las diferentes áreas disciplinares. Estado de avance: Idem anterior

4. FORMACIÓN DE RECURSOS HUMANOS

Lo lideran Directora, Co directora, tres Investigadores Responsables y dos Directores de proyectos Formar. Actualmente hay en curso: cinco tesinas de grado de Ingeniería en Computación y dos estudiantes de Doctorado. Egresaron dos Ingenieros Biomédicos. Lo integran dieciocho investigadores, en los roles de: Director, Responsable, En Formación o Colaborador.

5. BIBLIOGRAFÍA

1.Wietse de Vries, Arenas, Muñoz y Saldaña. 2011. ¿Desertores o decepcionados?, distintas causas para abandonar los estudios universitarios.

2.Zabalza M, 2013. VIII Congreso Iberoamericano de Docencia Universitaria y Nivel Superior. VIII CIDU http://www.iberoamericano2014.unr.edu.ar/

3.Britos, Díaz, Morales, Vargas. 2016. "Los MOOC como propuesta para la estandarización de la calidad educativa".

4.Wikipedia. National research and education network. Wikipedia web site online in http://en.wikipedia.org/wiki/National_research_and_education_network. Updated 2015.

5. Francois Vallaeys.2011. La responsabilidad social universitaria: Un nuevo modelo universitario contra la mercantilización.

6. Rizzardini, Gütl, Chang, Morales. 2014. MOOC in latin america: Implementation and lessons learned. The 2nd International Workshop on Learning Technology for Education in Cloud Springer Proceedings in Complexity.:158.

7. VEDUCA. http://www3.veduca.com.br/. Updated 2015.

8. MOOC.es. http://mooc.es. Updated 2015.

9. MIT OpenCourseWare. 2005. Program evaluation findings report.

10. edX. We're empowering learning in the classroom and around the globe. www.edx.org. Web site. https://www.edx.org/about-us.

11. Ibl studios education. Annotation tools inside open edX introduce a new paradigm in online learning. http://iblstudios.com/annotation-tools- inside-open-edx-introduce-a-new-paradigm- inonline-learning/.

12. Barchini, 2005. G. Métodos "I+D" de la Informática. Universidad Nacional de Santiago del Estero, Argentina. 2005.

13. xConsortium. OpenedX. code.edx.org Web site http://code.edx.org/.

14. Jordi Adell, 2017. "Las Tic no hacen bueno o malo a un Docente".

15. Edutec-e, 2018. www.edutec.es/congresos/xxi-congreso-internacional-edutec.

16. Kuna, García Martínez, Villatoro. 2010. "Identificación de causales de abandono de estudios universitarios. Uso de procesos de explotación de información". Revista Iberoamericana de Tecnología en Educación y Educación en Tecnología, TE&ET.

17. Formia, Lanzarini, Hasperué. 2013. Caracterización de la deserción universitaria en la UNRN utilizando Minería de Datos. Un caso de estudio. Revista Iberoamericana de Educación en Tecnología y Tecnología en Educación N°11. FI, UNLP, La Plata.

18. Díaz, Martins, García Martinez: 2015. "Descubrimiento De Patrones Socio-Económicos De Población Estudiantil De Carreras De Ingeniería Basado En Tecnologías De Explotación De Información". En Memorias de Congreso TE&ET.

19. Dehnadi y Bornat. 2006. Thecamel has twohumps (workingtitle). School of Computing. MiddleessexUniversity, UK.

20. Bornat; Dehnadi y Hamilton. 2008. Mentalmodels, Consistency, and Programming Aptitude. Autralian Computer Society. ACE.

21. Britos, Arias, Hirschfeld, 2015. Los MOOC un desafío para Latinoamérica

22. John Swope. A comparison of five free MOOC platforms for educators. EdTech Web site. http://www.edtechmagazine.com/higher/article/2014/02/comparison-five-free-mooc-platformseducators2015.

23. David Murphy's. The open source LMS battle hots up. EduBlog Web site. http://opob.edublogs.org/2013/05/02/the-open- source-lms-battle-hots-up/2015.

Uso de las Tecnologías de la Información y de la Comunicación para la Construcción de Espacios Institucionales

Julieta Gatica, Camila Olguín, José Montejano, Mariano Luzza, Mario Berón, Analía Zaldua, Germán Montejano, Ricardo Tertusio
Universidad Nacional de San Luis
Facultad de Ciencias Físico Matemáticas y Naturales
Área de Programación y Metodologías de Desarrollo de Software
Correo Electrónico: {mluzza,mberon,gmonte}@unsl.edu.ar
{jag81295,camimt25,jose.p.montejano,anazaldua,ricardodanielt}@gmail.com

Resumen

La educación es uno de los principales pilares sobre los que se cimienta el crecimiento social. En muchas situaciones los distintos contextos sociales y geográficos limitan a los miembros de una comunidad de recibir una educación de calidad. Es por esto que la elaboración de estrategias que permitan romper barreras de cualquier índole, que dificulten la educación, es una actividad de valor para todo tipo de sociedad.
Una de las formas de paliar los inconvenientes mencionados en el párrafo precedente consiste en el uso de las Tecnologías de la Información y de la Comunicación para brindar contenidos a distancia. En la actualidad, la Internet y la posibilidad de adquirir aparatos tecnológicos que se conectan a ella permite disminuir abruptamente las barreras existentes entre las sociedades situadas en distintas regiones. Esta peculiaridad es muy importante para diseñar soluciones que propendan a una inclusión educativa de calidad.
En este artículo se presenta una línea de investigación que aborda el diseño y construcción de espacios institucionales para evitar que la internet y tecnologías asociadas estén subutilizadas por la falta de diseño de un software que permita la integración de las actividades educativas y así solucionar el problema de conexión antes mencionado.

Palabras clave

Instituciones conectadas - Vinculación educativa - Repositorio institucional - Gestor de sitios web institucionales.

Contexto

La presente línea de investigación se enmarca en el Proyecto Bilateral (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 de Investigación (P-031516): "Ingeniería de Software: Conceptos, Prácticas y Herramientas para el Desarrollo de Software de 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

Hoy en día el uso de las tecnologías de la información y de la comunicación (TICs) en las instituciones educativas es una realidad insoslayable. Muchas campañas gubernamentales están orientadas a proporcionar, a los distintos miembros de la comunidad educativa, computadoras y tablets con el propósito de fomentar el uso de las TICs en educación y de incrementar las posibilidades educativas de la sociedad. Concomitantemente con este esfuerzo, el acceso a la tecnología por parte de los ciudadanos es más factible (en la mayoría de los hogares se puede encontrar al menos una computadora, smartphone, tablet, etc.). Esta característica hace que las estrategias educativas puedan salir del ámbito escolar o institucional y trasladarse a cualquier parte del planeta.

Lo mencionado en el párrafo precedente es una ventaja que los avances tecnológicos proveen. Los mismos, en la mayoría de los casos, no son explotados como se debe. Es común asistir a las instituciones escolares y observar la existencia de computadoras, netbooks, tablets, conexiones a internet, etc. las cuales en el mejor de los casos se utilizan para buscar información para resolver prácticos de las materias que cursan. Sin lugar a dudas, lo previamente mencionado es una ventaja respecto de épocas pasadas donde tal tipo de actividades eran impensadas. Sin embargo, muchas de las posibilidades que proporcionan las TICs no son utilizadas. A modo de ejemplo, se puede mencionar la poca vinculación que existe entre las escuelas y las instituciones de nivel superior e incluso con aquellas del mismo nivel. Esto se debe a que es necesario construir aplicaciones que posibiliten dicha interacción. Para alcanzar el objetivo antes mencionado, se debe contar con docentes o profesionales capacitados en el desarrollo de software. Si bien las instituciones educativas cuentan con profesionales con ese perfil, los mismos están dedicados a tareas docentes y generalmente no disponen del tiempo y recursos suficientes para poder llevar adelante un proyecto de software que provea una solución al problema de integración educativa.

Es importante mencionar que lo descrito en los párrafos anteriores no es un mero enunciado teórico sino más bien una realidad concreta observada y vivida por los docentes en su quehacer educativo.

2. Líneas de investigación y desarrollo

Las TICs se definen como “un conjunto diverso de herramientas y recursos tecnológicos usados para comunicar, crear, diseminar, almacenar y administrar información”[1]. Cuando se usan apropiadamente, ayudan a expandir el acceso a la educación, a elevar la calidad educativa, a fortalecer su importancia en el lugar de trabajo, etc., posibilitando que tanto la enseñanza como el aprendizaje se conviertan en un proceso activo conectado a la vida real. Sin embargo, la integración de las TICs es un proceso complejo, ya que no solo requiere de las tecnologías, sino también de pedagogía, preparación institucional, competencia docente, entre otras cosas.

En la actualidad, existen numerosas investigaciones acerca de la mejor forma para aplicar las TICs en contextos educativos. A través de una encuesta realizada a docentes de ingeniería eléctrica en la Universidad de Bechar, Algeria, se obtuvo que los docentes que hacen uso de las TICs en su proceso de enseñanza notaron que estas apoyan la

participación del alumno, crean un marco de justificación, facilitan el trabajo colaborativo y aumentan los factores de autonomía y creatividad en el alumno [2].

Otros estudios afirman que el avance de las TICs ha permitido a muchos países del mundo avanzar y crear una ventaja competitiva, y consideran que su implementación en la educación se considera el papel más crítico en la creación de una sociedad competitiva [3].

Teniendo en cuenta lo antes mencionado, se puede pensar que el problema principal en cuanto al uso de las TICs en la educación es la falta de un software que facilite la interacción hombre-máquina, y que ayude tanto a los profesores como a los estudiantes en el proceso de aprendizaje/enseñanza.

La evolución de las TICs, trajo consigo lo que se conoce como Sistema de Administración de Contenidos (CMS por sus siglas en inglés). Hoy en día, los CMS se utilizan en abundancia ya que son una plataforma de software que soporta aplicaciones basadas en web, de una manera fácil de usar. En contextos educativos, la explotación de estas plataformas ofrece soluciones integradas en la distribución del material del curso, gestión de los estudiantes e interacción entre las partes interesadas.

Los enfoques que utilizan CMSs en el campo educativo han sido enormemente estudiados. Autores como Segeč y Kubina [4], establecen que, teniendo cuidado en el análisis y la planificación de las características, y respetando las necesidades específicas de la comunidad educativa, se pueden crear sitios web para instituciones con mucha facilidad.

Por otro lado, Krouska et al. [5], presentaron una evaluación de los CMSs en la educación, a través de la cual se concluyó que ofrecen una variedad de opciones de personalización, tienen una gran cantidad de módulos disponibles que se pueden instalar para ampliar las capacidades del sistema, y se pueden desarrollar sitios web profesionales con mucha facilidad.

Muchos sitios, a menudo, son incapaces de recuperar rápidamente el ritmo cambiante de la gran cantidad de información. Cuando se amplía el sitio web, la integración de toda la información se vuelve más compleja; a veces hasta lleva a una reconstrucción del sitio. En tal circunstancia, Dang et al. [6] proponen la utilización de un CMS para gestionar y facilitar el sistema de contenido digital. Así, construir un sitio web se logra separando la administración de contenidos (base de datos o documentos independientes) del diseño (plantillas).

Cuando se habla de contenido, en esta línea de investigación, se hace referencia a lo que se considera repositorio institucional.

Los repositorios institucionales han alcanzado un nivel significativo de diversificación y desarrollo, y ahora se han convertido en un importante componente en la educación superior debido a los procesos de creación, publicación y distribución de contenido, aumentando su potencial e impacto. Al proporcionar a los alumnos de los recursos educativos pertinentes, se mejoran las condiciones de enseñanza y aprendizaje [7].

Las razones por las que se desarrollan tecnologías relacionadas a los repositorios institucionales es que hacen que la investigación científica y otra información académica producida, puedan ser integrales y estén centralizados a largo plazo. Además evita la existencia de información académica dispersa y garantiza el uso efectivo del conocimiento.

Al mismo tiempo, esto puede mejorar el status y el valor académico de la universidad, ya que permite un acceso abierto global mejorando los canales de distribución así como también, el medio ambiente [8].

Otra razón para la creación de un repositorio es la posible reutilización de los recursos educativos. Esto se ha descrito como Economía de Objetos de Aprendizaje [9], donde el material se comparte, luego se reutiliza y finalmente se mejora.

A pesar de la fuerza de esta base teórica, los profesores aún no han adoptado estas prácticas de reutilización. Petrides [10] realizó un análisis en el que estudia el alcance y la naturaleza de las prácticas de reutilización, y aquellas prácticas que involucran Recursos Educativos Abiertos (REA). Las siete categorías principales de reutilización fueron: 1) cambios visuales y técnicos, 2) edición general, 3) cambios relacionados con la colaboración, 4) cambios en los metadatos, 5) modularización, 6) traducciones de idiomas y 7) otros comportamientos de reutilización.

Finalmente, teniendo en cuenta que para una efectiva y eficiente experiencia de los usuarios en el uso del software, debe existir una Interfaz Gráfica de Usuario (GUI por sus siglas en inglés) que sea amigable.

Uno de los muchos beneficios al utilizar un CMS es que estos permiten aportar un valor añadido a los usuarios que visitan al sitio, ya que proporcionan información de calidad sobre el contenido, ofrecen una mejor atención al cliente, etc.

Al mantener el sitio web actualizado y aportar información de calidad periódicamente, aumenta la probabilidad de que el visitante vuelva, permitiendo así fidelizar a los usuarios existentes y atraer nuevos usuarios.

3. Resultados obtenidos/esperados

A través de los trabajos realizados por los integrantes de esta línea de investigación se han podido obtener diferentes resultados, los cuales se mencionan a continuación:

- Vinculación 3.0 es una herramienta que tiene como objetivo facilitar la interacción entre las diferentes instituciones educativas y sus comunidades asociadas. Para tal fin, la aplicación se ejecuta en un servidor con la configuración necesaria para dar respuesta en un tiempo razonable a los requerimientos institucionales, y proporciona diferentes funcionalidades que son muy útiles para las instituciones y sus comunidades educativas, como lo son: asignación de espacio en disco, creación de manera automática de sus sitios web, un repositorio compartido de recursos educativos abiertos, etc. Es un sistema que responde a diversos retos como la escasez de recursos financieros, cambios tecnológicos, transformaciones de recursos educativos, etc.
- A través del repositorio institucional, esta herramienta puede convertirse en un instrumento eficaz de promoción de la universidad, al establecer una interacción de conocimiento mutuo entre el nivel medio y el nivel superior, con el fin de que la percepción de su imagen sea acorde con la realidad vivida por los estudiantes y por los distintos usuarios de la comunidad educativa de dicho repositorio.
- La experiencia de utilizar Vinculación 3.0 en un colegio fue muy satisfactoria dado que las funcionalidades del sistema resultaron cubrir varios requerimientos de las instituciones educativas según la opinión del equipo de conducción de la institución (usuarios reales de Vinculación 3.0).

Obviamente, Vinculación 3.0 puede seguir creciendo a través del perfeccionamiento de

las funcionalidades que posee y la incorporación de otras nuevas que irán surgiendo a medida que sea utilizado por más instituciones. Como trabajo futuro se pretende:

- Incorporar mayor contenido al repositorio, y obtener una mayor comunidad que comparta, reutilice, y mejore los recursos educativos abiertos.
- Actualizar las GUIs del sitio de acuerdo a las recomendaciones y el feedback obtenido por parte de los usuarios, para mejorar la amigabilidad y la gestión por parte de los administradores.

4. Formación 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 maestría 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 mencionados en el contexto del presente documento.

5. Bibliografía

[1]. Victoria L. Tinio. ICT in Education.

[2]. Mammar Belagra, Chellali Benachaiba, Boutkhil Guemid. “Using ICT in Higher Education: Teachers of Electrical Engineering Department at the University of Bechar: Case Study”. IEEE Global Engineering Education Conference (EDUCON), 17, 20 April 2012.

[3]. Hardika Dwi Hermawan, Danis Nurul Yunita, Nurhamsi Deswila. “Implementation of ICT in Education in Indonesia during 2004-2017”. International Symposium on Educational Technology (ISET), 31 July, 2 Aug, 2018.

[4]. P. Segeč and M. Kubina, "Using of cms systems for schools and education," In Proc. International Conference of Virtual University, p. 5, 2008.

[5]. Akrivi Krouska, Christos Troussas, Maria Virvou. “Comparing LMS and CMS Platforms Supporting Social e-Learning in Higher Education”. 8th International Conference on Information, Intelligence, Systems & Applications (IISA), 27, 30 Aug, 2017.

[6]. Dang Xiaochao, Ma Wei, Li Yan. “UNDP403 Distance Education Website Based on WEB2.0 CMS System”. IEEE International Symposium on IT in Medicine and Education, 12, 14 Dec, 2008.

[7]. Virginia Rodés-Paragarino, Adriana Gewerc Barujel, Martín Llamas. “Use of Repositories of Digital Educational Resources. State of the Art Review”. IEEE Revista Iberoamericana de Tecnologías del Aprendizaje, Vol. 11, 15 April 2016.

[8]. Li Naiwen, Zhao Xin. “The Study On Constructing Institutional Repository Of University”. Fourth International Conference on Multimedia Information Networking and Security, 2, 4 Nov, 2012.

[9]. C. Duncan, R. House. “Learning object economies: Barriers and drivers” ELearnInternational, vol. 18, p. 19, 2004.

[10]. L. Petrides, L. Nguyen, A. Kargliani, C. Jimes. “Open educational resources: Inquiring into author reuse behaviors”. Technologies across learning contexts, Springer, 2008, pp. 344–353.

Tesis

Doctoral

RESUMEN TESIS DOCTORAL

"MODELO DE PROCESO PARA ELICITACIÓN DE REQUERIMIENTOS EN PROYECTOS DE EXPLOTACIÓN DE INFORMACIÓN"

Autor: **María Florencia Pollo Cattaneo** (flo.pollo@gmail.com)

Doctorado en Ciencias Informáticas Facultad de Informática - Universidad Nacional de La Plata

Directores: **Lic. Patricia Pesado, Dra Paola Britos**

Fecha de exposición: 10 de Mayo de 2018

Resumen

El presente artículo resume el trabajo de la tesis doctoral "MODELO DE PROCESO PARA ELICITACIÓN DE REQUERIMIENTOS EN PROYECTOS DE EXPLOTACION DE INFORMACIÓN", cuyo propósito es desarrollar un proceso metodológico en el marco de los proyectos de Explotación de Información que brinde solución a la necesidad de gestionar en forma adecuada los requerimientos en este tipo de proyectos y de esta manera, promueva una mejora en el campo de la Ingeniería de Requerimientos. Se presenta un Modelo de Proceso para la Elicitación de Requerimientos en proyectos de Explotación de Información, el cual pone énfasis en la definición del proyecto, la educción y conceptualización del negocio, así como la identificación de los procesos de Explotación de Información a utilizar. Para ello, en primer lugar, se presenta la estructura del modelo propuesto, la cual es descripta a partir de sus fases y actividades; y, en segundo lugar se realiza la presentación de un conjunto de procesos de formalización de requisitos que son alimentados por las actividades del modelo antes mencionado. Además, se presentan dos casos de validación pertenecientes a diferentes dominios, de forma de mostrar que el modelo de proceso propuesto es válido para cualquier tipo de dominio.

Palabras clave: Explotación de Información, Requerimientos, Ingeniería de Requerimientos

1. INTRODUCCIÓN

La Inteligencia de Negocio propone un abordaje interdisciplinario cuya finalidad es generar conocimiento que contribuya con la toma de decisiones de gestión y con la generación de planes estratégicos en las organizaciones [Thomsen, 2003]. Para poder elaborar ese conocimiento, cuenta con una sub-disciplina de la Informática llamada Explotación de Información que aporta las herramientas de análisis y síntesis para extraer el conocimiento no trivial que se encuentra alojado, en forma implícita, en los datos disponibles de diferentes fuentes de información [Schiefer et al., 2004]. Es por esto, que las metodologías que se utilizan para el desarrollo de proyectos en Explotación de Información buscan organizar el proceso de descubrimiento de patrones en los almacenes de datos de una organización [Britos, 2008]. Estas metodologías consideran que la especificación de requerimientos es una de las actividades tempranas del proyecto. [Britos et al., 2008]. Varios autores [Langseth & Vivatrat, 2003; Grigori et al., 2004; Chung et al., 2005] señalaron la necesidad de mejorar las metodologías de Explotación de Información pero sus mejoras se orientaron a la definición de objetivos [Arnth-Jensen, 2006] y a la especificación de algunas tareas específicas tales como el análisis de datos o el hallazgo de patrones [Dasgupta & Vankayala, 2007], descuidando esta principal actividad que debe ser tenida en cuenta para que los proyectos puedan concluir con éxito.

Por lo tanto, se detecta la necesidad de proveer a la comunidad perteneciente a la Explotación de Información un conjunto de herramientas que permitan realizar una correcta administración de los requerimientos en este tipo de proyectos.

En este contexto, se ha propuesto como objetivo general de este trabajo de tesis desarrollar un proceso metodológico en el marco de los proyectos de Explotación de Información que brinde solución a la necesidad de gestionar en forma adecuada los requerimientos en este tipo de proyectos y de esta manera, promueva una mejora en el campo de la Ingeniería de Requerimientos.

La tesis se enfoca en plantear un modelo de proceso para la elicitación de requerimientos definiendo un conjunto de fases, actividades asociadas a estas fases y tareas que se realizan en cada actividad, de modo de facilitar a los ingenieros de Explotación de Información, la administración de los requerimientos. Para proponer este modelo se profundiza en el Modelo de Proceso para Proyectos de Explotación de Información definido en [Vanrell et al., 2010; 2012], el cual se basa en la metodología CRISP-DM [Chapman et al., 2000] e incorpora principios de la metodología COMPETISOFT [Oktaba et al., 2007] que considera los aspectos de las Pequeñas y Medianas Empresas (PyMEs).

La tesis formula contribuciones sobre: actividades de gestión de requerimientos en proyectos de Explotación de Información y procesos de documentación de requerimientos.

2. ESTADO DE LA CUESTIÓN

La Explotación de Información consiste en la aplicación de herramientas de análisis y síntesis con el objetivo de extraer conocimiento no trivial que se encuentra distribuido en forma implícita en los datos disponibles de diferentes fuentes de información dentro de una organización [Schiefer et al., 2004]. Este conocimiento es desconocido previamente y puede ser útil para la toma de decisiones en una organización [Thomsen, 2003]. Para un experto, o para el responsable de un sistema de información, generalmente no son los datos en sí lo más relevante, sino el conocimiento que se encierra en sus relaciones, fluctuaciones y dependencias [Curtis et al., 1992]. Si este conjunto de relaciones refleja la realidad, por lo cual son válidas, aportan algo novedoso y útil para la toma de decisiones [Kanungo, 2005].

Una vez definido el concepto de Explotación de Información es necesario enfocarse en las metodologías existentes que se aplican en la ejecución de este tipo de proyectos, que poseen madurez respecto al desarrollo de un proyecto de este tipo, entre las que se destacan CRISP-DM [Chapman et al., 2000], P3TQ [Pyle, 2003] y SEMMA [SAS, 2008]. En este trabajo sólo se considera la metodología CRISP-DM por ser considerada la que mejor maneja las actividades correspondientes a la gestión de estos proyectos. Por otro lado, también se profundiza sobre el Modelo de Proceso para Proyectos de Explotación de Información. Dicho modelo de proceso se basa en la metodología CRISP-DM brindando solución a varias de sus limitaciones y puntos débiles.

Los proyectos de Explotación de Información no escapan de la necesidad de contar con requerimientos claros, completos y estables para poder completar el proyecto en forma exitosa [Pollo-Cattaneo et al., 2012]. Sin embargo, son diferentes los requerimientos de este tipo de proyecto, dado que un proyecto de Explotación de Información no busca la construcción de un producto software, como en el caso de la Ingeniería de Software. Es por esto que los requerimientos no están enfocados a definir las funcionalidades y restricciones que deberá cumplir un sistema software.

Al comienzo de todo proyecto de Explotación de Información se deben identificar los objetivos del proyecto que describen la necesidad general del cliente, es decir lo que el cliente espera obtener como

resultado final del proyecto [Britos et al., 2008]. Este tipo de objetivos normalmente se encuentran muy relacionados con las metas estratégicas y tácticas del negocio. Para poder lograr un mejor entendimiento sobre los aspectos del proyecto, se tiene la dificultad adicional de no manejar correctamente el vocabulario del negocio utilizado por los miembros de la organización.

Una vez que estos objetivos son identificados es necesario realizar un reconocimiento inicial de las fuentes de información disponible en la organización identificando cuáles fuentes se encuentran informatizadas (en repositorios de datos) y cuáles no. En el caso de los repositorios es muy importante también determinar la estructura y naturaleza de los datos disponibles.

A partir del análisis de los objetivos del proyecto y los repositorios de datos, es posible delimitar el alcance del proyecto en un conjunto de objetivos de requerimientos. Luego éstos podrán ser resueltos a través de la aplicación de procesos de Explotación de Información aplicando los algoritmos de Minería de Datos correspondientes [García-Martínez et al., 2011b]. De esta forma se podrá dar solución al problema de negocio que dio origen al proyecto.

En [Britos et al., 2008] se señala la necesidad de adaptar el proceso tradicional de especificación de requerimientos de sistemas software para proyectos de Explotación de Información y se propone un proceso de educción de requisitos que contempla la elaboración de un conjunto de plantillas, que son utilizadas para documentar los conceptos educidos durante las fases de comprensión del negocio y del dominio de proyecto. Cada plantilla se asocia a un concepto educido y contiene una descripción detallada del mismo, permitiendo la evolución del concepto a lo largo del proyecto. De esta forma, a partir de la información contenida en las plantillas, el ingeniero de Explotación de Información puede realizar las actividades del proyecto, tal como identificar los repositorios de datos, determinar el modelo y seleccionar las herramientas a utilizar dentro del proyecto.

En relación a la documentación de Requisitos en Proyectos de Explotación de Información, en [Britos et al., 2008] se señala la necesidad de adaptar el proceso tradicional de especificación de requerimientos de sistemas software para proyectos de Explotación de Información y se propone un proceso de educción de requisitos que contempla la elaboración de un conjunto de plantillas, que son utilizadas para documentar los conceptos educidos durante las fases de comprensión del negocio y del dominio de proyecto. Cada plantilla se asocia a un concepto educido y contiene una descripción detallada del mismo, permitiendo la evolución del concepto a lo largo del proyecto. De esta forma, a partir de la información contenida en las plantillas, el ingeniero de Explotación de Información puede realizar las actividades del proyecto, tal como identificar los repositorios de datos, determinar el modelo y seleccionar las herramientas a utilizar dentro del proyecto.

3. SOLUCIÓN

En virtud de lo expuesto, el trabajo de tesis tiene como objetivo proponer un Modelo de Proceso para la Elicitación de Requerimientos en proyectos de Explotación de Información, que busca brindar solución a la necesidad de gestionar los requerimientos en este tipo de proyectos, dado que las metodologías existentes hacen hincapié en la identificación de los datos disponibles sobre los cuales realizar los procesos de Explotación de Información pero dejan de lado las actividades asociadas a los requerimientos de los clientes e interesados.

El modelo propuesto, que se muestra en la figura 1, se estructura en las siguientes cinco fases:

1. Una primera fase de Definición del Proyecto, cuyo objetivo es definir el alcance del proyecto, los interesados y los objetivos que se deben alcanzar.

2. Una segunda fase de Educción de Procesos de Negocio, cuyo objetivo es identificar y relevar los procesos de negocio significativos al proyecto.

3. Una tercera fase de Educción de Datos de Procesos de Negocio, cuyo objetivo es identificar los repositorios de datos donde se almacena la información de los diferentes procesos de negocio y relevar los datos existentes en esos repositorios.

4. Una cuarta fase de Conceptualización del Negocio, cuyo objetivo es definir el negocio en términos de conceptos utilizados y vocabulario, para comprender el idioma utilizado en el negocio, descubrir las palabras específicas del mismo y cuál es el significado que el negocio le da a esas palabras específicas.

5. Una quinta fase de Especificación de Procesos de Explotación de Información, cuyo objetivo es identificar los procesos de Explotación de Información que se pueden utilizar para resolver los problemas identificados en el proceso de negocio y luego realizar la planificación del resto de las actividades del proyecto.

Figura 1. Fases del Modelo de Proceso Propuesto

A continuación se detallan en la tabla 1 las entradas y las salidas de la Fase de Definición del Proyecto.

Tabla 1. Entradas y Salidas de la Fase de Definición del Proyecto.

Fase	Actividad	Productos de Entrada	Procedimiento	Productos de Salida
Definición del Proyecto	Identificar los Objetivos del Proyecto	Definición Conceptual del Proyecto	Análisis de Definición Conceptual del Proyecto Entrevistas	Minuta de Reunión Conceptual Acta de Reunión de Inicio del Proyecto
				Plantilla de Objetivos del Proyecto Plantilla de Objetivos del Requisito
	Identificar los Interesados del Proyecto	Acta de Reunión de Inicio del Proyecto	Entrevistas Workshops	Plantilla de Interesados del Proyecto Plantilla de Criterios de Éxito del Proyecto Plantilla de Expectativas del Proyecto
	Identificar el Alcance del Proyecto	Plantilla de Interesados del Proyecto Plantilla de Criterios de Éxito del Proyecto Plantilla de Criterios de Éxito del Requisito	Análisis de Información recopilada Entrevistas	Definición de Alcance de Proyecto
				Plantilla de Restricciones del Proyecto Plantilla de Supuestos del Proyecto Plantilla de Restricciones del Requisito Plantilla de Supuestos del Requisito Plantilla de Riesgos del Proyecto Plantilla de Planes de Contingencia del Proyecto Plantilla de Riesgos del Requisito Plantilla de Planes de Contingencia del Requisito

En esta Fase, los Procesos de formalización utilizados son: Proceso de Identificación de Objetivos, Criterios de Éxito y Expectativas del Proyecto, Proceso de Identificación de Suposiciones y

Restricciones del Proyecto y Proceso de Identificación de Riesgos y Planes de Contingencia del Proyecto

A continuación, la tabla 2 resume las entradas y las salidas de la Fase de Educción de Procesos de Negocio.

Tabla 2. Entradas y Salidas de la Fase de Educción de Procesos de Negocio.

Fase	Actividad	Productos de Entrada	Procedimiento	Productos de Salida
Educción de Procesos de Negocio	Identificar Procesos de Negocio	Plantilla de Definición de Alcance de Proyecto Plantilla de Objetivos del Proyecto Plantilla de Objetivos del Requisito Plantilla de Criterios de Éxito del Proyecto Plantilla de Expectativas del Proyecto	Análisis de Documentación	Plantilla de Diagrama de Procesos de Negocio
	Relevar Procesos de Negocio	Plantilla de Diagrama de Procesos de Negocio	Entrevistas Workshops	Plantilla de Procesos de Negocio

La tabla 3, resume las entradas y las salidas de la Fase de Educción de Datos de Procesos de Negocio.

Tabla 3. Entradas y Salidas de la Fase de Educción de Datos de Procesos de Negocio.

Fase	Actividad	Productos de Entrada	Procedimiento	Productos de Salida
Educción de Datos de Procesos de Negocio	Identificar Repositorios de Datos	Plantilla de Diagrama de Procesos de Negocio Plantilla de Procesos de Negocio Documentación Escrita proveniente de Entrevistas realizadas a los Interesados del Negocio	Análisis de Documentación	Plantilla de Repositorios de Datos
	Relevar Datos del Negocio	Plantilla de Repositorios de Datos	Entrevistas Estructuradas	Plantilla de Estructura de Datos

La tabla 4, resume las entradas y las salidas de la Fase de Conceptualización del Negocio.

Tabla 4. Entradas y Salidas de la Fase de Conceptualización del Negocio.

Fase	Actividad	Productos de Entrada	Procedimiento	Productos de Salida
Conceptualización del Negocio	Construir el Diccionario del Negocio	Plantilla de Diagrama de Procesos de Negocio Plantilla de Procesos de Negocio Plantilla de Repositorios de Datos Plantilla de Estructura de Datos	Análisis de Documentación Entrevistas Estructuradas	Diccionario de Negocio
				Plantilla de Definiciones, Acrónimos y Abreviaturas Grafo de Relaciones entre Conceptos Plantilla de Atributos relacionados con los requisitos
	Construir el Modelo de Negocio	Plantilla de Repositorios de Datos Plantilla de Estructura de Datos Diccionario de Negocio	Análisis de Documentación	Diagrama de Modelo del Negocio

En esta Fase, el Proceso de formalización utilizado es: Proceso de Formalización de Dominios de Negocio

Finalmente, la tabla 5 resume las entradas y las salidas de la Fase de Especificación de Procesos de Explotación de Información.

Tabla 5. Entradas y Salidas de la Fase de Especificación de Procesos de Explotación de Información.

Fase	Actividad	Productos de Entrada	Procedimiento	Productos de Salida
Especificación de Procesos de EdI	Formalizar los Problemas de Negocio	Diagrama de Modelo de Negocio Definición de Alcance de Proyecto Plantilla de Objetivos del Proyecto Plantilla de Objetivos del Requisito	Análisis de Documentación	Plantilla de Problemas de Negocio Plantilla de Objetivos del Requisito
	Identificar los Procesos de EdI	Plantilla de Problemas de Negocio	Análisis de Documentación	Plantilla de Procesos de Explotación de Información Plantilla de Suposiciones del Requisito Plantilla de Restricciones del Requisito Plantilla de Atributos relacionados con los Requisitos
	Construir el Plan del Proyecto	Plantilla de Objetivos del Proyecto Plantilla de Objetivos del Requisito Plantilla de Interesados del Proyecto Plantilla de Riesgos del Proyecto Plantilla de Riesgos del Requisito Definición de Alcance de Proyecto Plantilla de Procesos de Explotación de Información	Análisis de Documentación	Plantilla de Plan del Proyecto

En esta Fase el Proceso de formalización utilizado es: Proceso de Revisión de Requisitos del Proyecto

Por otro lado, en relación con la documentación, se presentan cinco procesos de formalización de requisitos que son utilizados dentro del proceso de educción propuesto. De esta manera, los procesos de formalización se emplean para completar las plantillas propuestas en [Britos et al., 2008], las cuales son utilizadas para documentar los conceptos que fueron educidos en las diferentes fases del proceso de educción. Cada proceso aplica técnicas de modelado provenientes de la Ingeniería del Conocimiento. Los procesos que se describen son los siguientes: "Proceso de Identificación de Objetivos, Criterios de Éxito y Expectativas del Proyecto", "Proceso de Identificación de Suposiciones y Restricciones del Proyecto", "Proceso de Identificación de Riesgos y Planes de Contingencia del Proyecto", "Proceso de Formalización de Dominios de Negocio" y, "Proceso de Revisión de Requisitos del Proyecto".

4. VALIDACIÓN

Se presentan dos casos de validación pertenecientes a diferentes dominios, de forma de mostrar que el modelo de proceso propuesto es apto para cualquier tipo de dominio. El primer caso corresponde a una empresa metalúrgica que desea poseer mayor conocimiento sobre el consumo de recursos de cada actividad que se realiza dentro del área de producción, descubriendo patrones de comportamiento que permitan describir la utilización de los recursos en el área de producción de

acoplados . Estos patrones permitirán asistir a la toma de decisiones al momento de la planificación de la línea de ensamble de unidades

Luego, se presenta un segundo caso correspondiente a una empresa de turismo que desea mejorar su oferta de paquetes turísticos, que desea mejorar su oferta de paquetes turísticos teniendo en cuenta la proyección de gastos de sus clientes.

5. CONCLUSIONES Y FUTURAS LÍNEAS

El trabajo de tesis aporta a la comunidad, elementos para poder realizar una gestión integral de los requerimientos dentro del desarrollo de proyectos de Explotación de Información, pudiendo de esta manera, promover una mejora a las metodologías existentes que descuidan este aspecto.

Teniendo en cuenta que la descripción del problema, se articuló a partir de preguntas de investigación, las respuestas a cada una de ellas constituyen las aportaciones de la misma. A continuación se presenta el Sumario de Investigación.

La primera pregunta del Sumario consistía en lo siguiente: ¿Es posible definir un modelo de proceso en el marco de los proyectos de Explotación de Información que permita gestionar y documentar en forma eficaz los requerimientos de los clientes con la finalidad de que los proyectos finalicen en forma exitosa?

Para dar respuesta a este interrogante, se ha propuesto un modelo de proceso para proyectos de Explotación de Información que permite realizar una gestión global de los requerimientos

Como se ha podido construir este modelo, es posible responder las tres preguntas adicionales:

- Pregunta 1a: "¿Cuáles son las fases que incluye este modelo de proceso?"

El modelo de proceso contempla cinco fases, en las cuales se pone énfasis en la definición del proyecto, la educción y conceptualización del negocio y en la identificación de los procesos de Explotación de Información a utilizar. Las fases son las siguientes: Definición del Proyecto, Educción de Procesos de Negocio, Educción de Datos de Procesos de Negocio, Conceptualización del Negocio y Especificación de Procesos de Explotación de Información.

- Pregunta 2a: "¿Cuáles son las actividades vinculadas a cada fase?"

Las actividades que se definieron dentro del proceso son:

1. Para la fase de Definición del Proyecto, se han propuesto las siguientes actividades: [i] Identificar los Objetivos del Proyecto, [ii] Identificar los Interesados del Proyecto e [iii] Identificar el Alcance del Proyecto.

2. Para la fase de Educción de Procesos de Negocio, se han propuesto las siguientes actividades: [i] Identificar Procesos de Negocio y [ii] Relevar Procesos de Negocio.

3. Para la fase de Educción de Datos de Procesos de Negocio, se han propuesto las siguientes actividades: [i] Identificar Repositorios de Datos y [ii] Relevar Datos del Negocio.

4. Para la fase de Conceptualización del Negocio, se han propuesto las siguientes actividades: [i] Construir el Diccionario del Negocio y [ii] Construir el Modelo de Negocio.

5. Para la fase de Especificación de Procesos de Explotación de Información, se han propuesto las siguientes actividades: [i] Formalizar los Problemas de Negocio, [ii] Identificar los Procesos de Explotación de Información y [iii] Construir el Plan del Proyecto.

- Pregunta 3a: "¿Cómo se puede realizar una adecuada documentación de los requerimientos utilizando las plantillas definidas?"

Para realizar una adecuada documentación de los requerimientos utilizando las plantillas definidas, se determinaron cinco procesos de formalización que se aplican dentro de las fases del proceso. Los procesos se detallan a continuación:

1. Proceso de Identificación de Objetivos, Criterios de Éxito y Expectativas del Proyecto, que es llamado por las actividades "Identificar los Objetivos del Proyecto" e "Identificar los Interesados del Proyecto" de la fase de "Definición del Proyecto" y se utiliza para realizar la documentación de los objetivos, criterios de éxito y expectativas del proyecto.

2. Proceso de Identificación de Suposiciones y Restricciones del Proyecto, que es llamado por la actividad "Identificar el Alcance del Proyecto" de la fase de "Definición del Proyecto" y se utiliza para realizar la documentación de las restricciones y suposiciones del proyecto.

3. Proceso de Identificación de Riesgos y Planes de Contingencia del Proyecto, que es llamado por la actividad "Identificar el Alcance del Proyecto" de la fase de "Definición del Proyecto" y tiene como objetivo documentar los riesgos del proyecto y los planes de contingencia asociados a cada riesgo.

4. Proceso de Formalización de Dominios de Negocio, que es llamado por la fase "Conceptualización del Negocio" y se utiliza para realizar la documentación del vocabulario utilizado dentro del dominio y para establecer las relaciones entre los conceptos educidos del negocio.

5. Proceso de Revisión de Requisitos del Proyecto, que es llamado por las actividades "Formalizar los Problemas de Negocio" e "Identificar los Procesos de Explotación de Información" de la fase de "Especificación de Procesos de Explotación de Información" y se lo utiliza para realizar la verificación de los objetivos del requisito, las restricciones y las suposiciones relacionadas a estos objetivos.

El modelo de proceso propuesto (incluyendo fases, actividades, tareas y procesos de formalización asociados) ha sido validado en tres dominios diferentes: el primer caso que tuvo como objetivo la detección de daños y averías producidas en vehículos cero kilómetros, a lo largo del circuito logístico; el segundo caso correspondiente a una empresa metalúrgica que deseaba poseer mayor conocimiento sobre el consumo de recursos de cada actividad realizada dentro del área de producción y el tercer caso, correspondiente a una empresa de turismo que deseaba mejorar su oferta de paquetes turísticos.

Durante el desarrollo de esta tesis han surgido cuestiones que, si bien no son centrales al tema abordado en la misma, constituyen temas concomitantes que (en opinión del tesista) darían lugar a las siguientes líneas de investigación futuras:

I. Si bien el proceso propuesto en esta tesis aporta sistematicidad a la gestión integral de requerimientos en proyectos de Explotación de Información y permite realizar la documentación de los requerimientos utilizando las plantillas definidas, queda como tema de trabajo abierto, el poder detallar procedimientos estructurados con el paso a paso para cada proceso de formalización, de forma de que las técnicas de cada uno de estos procesos sean más automatizables de realizar por un ingeniero en Explotación de Información.

II. Aunque los proyectos que fueron utilizados en la validación del proceso propuesto se consideran representativos, se estima recomendable aplicar el proceso en una mayor cantidad de proyectos, tanto dentro de empresas PyMEs como en empresas de mayor envergadura de manera de evaluar el proceso bajo situaciones de negocio nuevas y a partir de esta evaluación, poder realizar una generalización del proceso propuesto para todo tipo de empresas.

III. A los efectos de ayudar a los integrantes del equipo de desarrollo en el seguimiento de las plantillas y actas de reunión que propone el proceso, se considera conveniente desarrollar una herramienta software de apoyo al proceso propuesto con el fin de identificar el estado de las plantillas y actas propuestas por el proceso, distinguir sus diferentes versiones, informar lo que contiene cada versión e incluir un histórico de cambios.

6. REFERENCIAS

Arnth-Jensen, N. 2006. Applied Data Mining for Business Intelligence. Tesis de Maestría. Informatics and Mathematical Modelling Department, Technical University of Denmark.

Britos, P., Fernandez, E., García-Martínez, R. 2007. Propuesta Metodológica para la Educción de Requisitos en Proyectos de Explotación de Información. IX Workshop de Investigadores en Ciencias de la Computación, pp. 497-501. ISBN 978-950-763-075-0.

Britos, P. 2008. Procesos de Explotación de Información basados en Sistemas Inteligentes. Tesis Doctoral. Universidad Nacional de La Plata, Facultad de Informática. La Plata. Argentina. Disponible en: http://www.iidia.com.ar/rgm/tesistas/td-pb-fi-unlp.pdf. Último Acceso: Enero de 2016.

Britos, P., Dieste, O., García-Martínez, R. 2008. Requirements Elicitation in Data Mining for Business Intelligence Projects. Advances in Information Systems Research, Education and Practice. David Avison, George M. Kasper, Barbara Pernici, Isabel Ramos, Dewald Roode Eds. (Boston: Springer). IFIP International Federation for Information, 274: 139–150.

Britos, P. & García-Martínez, R. 2009. Propuesta de Procesos de Explotación de Información. Proceedings XV Congreso Argentino de Ciencias de la Computación. Workshop de Base de Datos y Minería de Datos, pp. 1041-1050. ISBN 978-897-24068-4-1.

Chapman, P., Clinton, J., Keber, R., Khabaza, T., Reinartz, T., Shearer, C. y Wirth, R. 2000. CRISP-DM 1.0 Step by step BI guide. Edited by SPSS. Disponible en: http://tinyurl.com/crispDM. Último Acceso: Enero de 2016.

Chung, W., Chen, H., Nunamaker, J. 2005. A Visual Framework for Knowledge Discovery on the Web: An Empirical Study of Business Intelligence Exploration. Journal of Management Information Systems, 21(4): 57-84.

Dasgupta, S. & Vankayala, K. 2007. Developing Realtime Business Intelligence Systems.The Agile Way. Proceedings 1st Annual 2007 IEEE Systems Conference. Pág, 63-69.

Garcia–Martinez, R., Britos, P., Pollo–Cattaneo, F., Rodriguez, D., Pytel, P. 2011b. Information Mining Processes Based on Intelligent Systems. Proceedings of II International Congress on Computer Science and Informatics (INFONOR–CHILE 2011). Pág. 87–94. ISBN 978–956–7701–03–2.

Grigori, D., Casati, F., Castellanos, M., Dayal, u., Sayal, M., Shan, M. 2004. Business Process Intelligence. Computers in Industry 53(3): 321-343. 5.

Kanungo, S. 2005. Using Process Theory to Analyze Direct and Indirect Value-Drivers of Information Systems. Proceedings of the 38th Annual Hawaii International Conference on System Sciences, pp. 231-240.

Langseth, J. & Vivatrat, N. 2003. Why Proactive Business Intelligence is a Hallmark of the Real-Time Enterprise: Outward Bound. Intelligent Enterprise 5(18): 34-41. 4.

Oktaba, H., Garcia, F., Piattini, M., Ruiz, F., Pino y F.J. y Alquicira, C. 2007. Software Process Improvement: The COMPETISOFT Project. Computer 40(10), pp. 21-28.

Pollo-Cattaneo, M., García–Martínez, R., Britos, P., Pesado, P., Bertone, R., Rodríguez, D., Merlino, H., Pytel, P., Vanrell, J. 2012. Elementos para una Ingeniería de Explotación de Información. Proyecciones 10(1): 67–84. ISSN 1667–8400.

Pollo-Cattaneo, M., Pytel, P., Vegega, C., Mansilla, D., Pesado, P., García–Martínez, R., Britos, P. 2013. Elementos para la Gestión de Requerimientos en Proyectos de Explotación de Información. Congreso Nacional de Ingeniería Informática / Sistemas de Información. ISSN: 2346-9927.

Pyle, D. 2003. Business Modeling and Business Intelligence. Morgan Kaufmann Publishers.

SAS. 2008. SAS Enterprise Miner: SEMMA. Disponible en: http://tinyurl.com/semmaSAS. Último Acceso: Enero de 2016.

Schiefer, J., Jeng, J., Kapoor, S. y Chowdhary, P. 2004. Process Information Factory: A Data Management Approach for Enhancing Business Process Intelligence. Proceedings 2004 IEEE International Conference on E-Commerce Technology, pp. 162-169.

Vanrell, J., Bertone, R. y García-Martínez, R. 2010. Modelo de Proceso de Operación para Proyectos de Explotación de Información. Anales del XVI Congreso Argentino de Ciencias de la Computación, pp. 674-682. ISBN 978-950-9474-49-9.

Vanrell, J., Bertone, R., García-Martínez, R. 2012. Un Modelo de Procesos para Proyectos de Explotación de Información. Proceedings Latin American Congress on Requirements Engineering and Software Testing, pp. 46-52. ISBN 978-958-46-0577-1.

Thomsen, E. 2003. BI's Promised Land. Intelligent Enterprise, 6(4), pp. 21-25.

Un Modelo de Arquitectura para un Sistema de Virtualización Distribuido

Autor: Pablo Andrés Pessolani
Directores:
Dr. Toni Cortés – UPC – BSC- Barcelona – España.
Dr. Fernando G. Tinetti – UNLP, CIC Prov. de Bs. As. – La Plata – Argentina
Co-Director:
Dr. Silvio Gonnet – INGAR (CONICET - UTN FRSF) – Santa Fe - Argentina.

Resumen de Tesis presentada para obtener el grado de Doctor en Ciencias Informáticas

Universidad Nacional de La Plata – Facultad de Informática

Defensa de Tesis: 21 de Noviembre de 2018 – La Plata
Video Online: https://www.youtube.com/watch?v=Ha9Cx4sWimk
Calificación obtenida: 10 (diez).

Palabras claves: Virtualización, Contenedores, Sistemas Distribuidos, IaaS.

RESUMEN

Si bien los Sistemas Operativos disponen de características de seguridad, protección, gestión de recursos, etc., éstas parecen ser insuficientes para satisfacer los requerimientos de los sistemas informáticos que suelen estar permanente y globalmente conectados. Las actuales tecnologías de virtualización han sido, y continúan siendo masivamente adoptadas para cubrir esas necesidades de sistemas y aplicaciones por sus características de particionado de recursos, aislamiento, capacidad de consolidación, seguridad, soporte de aplicaciones heredadas, facilidades de administración, etc. Una de sus restricciones es que el poder de cómputo de una Máquina Virtual (o un Contenedor) está acotado al poder de cómputo de la máquina física que la contiene. Esta tesis propone superar esta restricción abordando esta problemática con el enfoque de un sistema distribuido.

Para poder alcanzar mayores niveles de rendimiento y escalabilidad, los programadores de aplicaciones nativas para la Nube deben partirlas (factorizarlas) en diferentes componentes distribuyendo su ejecución en varias Máquinas Virtuales (o Contenedores). Dichos componentes se comunican mediante interfaces bien definidas tales como las interfaces de Web Services. Las Máquinas Virtuales (o Contenedores) deben configurarse, asegurarse y desplegarse para poder ejecutar la aplicación. Esto se debe, en parte, a que los diferentes componentes no comparten la misma instancia de Sistema Operativo por lo que no comparten los mismos recursos abstractos tales como colas de mensajes, mutexes, archivos, pipes, etc. El defecto de esta modalidad de desarrollo de aplicaciones es que impide una visión integral y generalizada de los recursos. En ella, el programador debe planificar la asignación de recursos a cada componente de su aplicación y, por lo tanto, no solo debe programar su aplicación sino también gestionar la distribución de esos recursos.

En este trabajo se propone un modelo de arquitectura para un Sistema de Virtualización Distribuido (DVS) que permite expandir los límites de un dominio de ejecución más allá de una máquina física, explotando el poder de cómputo de un cluster de computadores. En un DVS se combinan e integran tecnologías de Virtualización, de Sistemas Operativos y de Sistemas Distribuidos, donde cada una de ellas le aporta sus mejores características. Este modelo de arquitectura, por ejemplo, le brinda al programador una visión integrada de los recursos distribuidos

que dispone para su aplicación relevándolo de la responsabilidad de gestionarlos. El modelo de DVS propuesto dispone de aquellas características que son requeridas por los proveedores de servicios de infraestructura en la Nube, tales como: mayor rendimiento, disponibilidad, escalabilidad, elasticidad, capacidad de replicación y migración de procesos, balanceo de carga, entre otras.

Las aplicaciones heredadas pueden migrarse más fácilmente, dado que es posible disponer de la misma instancia de un Sistema Operativo Virtual en cada nodo del cluster de virtualización. Las aplicaciones desarrolladas bajo las nuevas metodologías para el diseño y desarrollo de software para la Nube también se benefician adaptándose su utilización a un sistema que es inherentemente distribuido.

INTRODUCCIÓN

Las actuales tecnologías de virtualización han sido y son masivamente adoptadas para satisfacer aquellos requerimientos en que los Sistemas Operativos (OS: Operating System) han presentado debilidades, tales como el aislamiento de fallos, de seguridad y de rendimiento. Además, la virtualización incorpora nuevas facilidades tales como el particionado de recursos, la posibilidad de realizar la consolidación de servidores, el soporte para la ejecución de aplicaciones legadas, la disponibilidad de herramientas que facilitan su gestión, etc. Todas estas características resultan muy atractivas para las organizaciones en general y para los proveedores de Servicios en la Nube (Cloud Services) en particular.

Actualmente, existen varias tecnologías de virtualización que permiten brindar Infraestructura como Servicio (IaaS: Infrastructure as a Service) desplegadas sobre un cluster de servidores enlazados por redes de alta velocidad. Las redes de almacenamiento (SAN: Storage Area Networks), los dispositivos de seguridad (firewalls de red, firewalls de aplicación, Sistemas de Detección y Prevención de Intrusos, etc.), los dispositivos de red (switches, routers, balanceadores de carga, etc.), y un conjunto de sistemas de gestión y monitoreo complementan la infraestructura (informática) requerida para brindar este tipo de servicios a terceros en Nubes Públicas, o en la propia compañía (on-premise) para el caso de Nubes Privadas.

Para establecer los límites de las responsabilidades entre el proveedor de Servicios en la Nube y el cliente, parece adecuado adoptar la definición del NIST respecto a Infraestructura como Servicio (IaaS):

> *La capacidad de proveer al consumidor procesamiento, almacenamiento, redes y otros recursos informáticos fundamentales sobre los cuales el consumidor puede implementar y ejecutar software arbitrario, pudiendo incluir sistemas operativos y aplicaciones. El consumidor no administra ni controla la infraestructura subyacente de la Nube, pero tiene control sobre los sistemas operativos, el almacenamiento y otras aplicaciones implementadas; y posiblemente un control limitado de los componentes de red seleccionados (por ejemplo, firewalls) [1].*

Las tecnologías de virtualización más utilizadas son *Virtualización de Hardware*, *Paravirtualización* y *Virtualización basada en Sistema Operativo*. Cada una de ellas presenta diferentes características en cuanto a niveles de consolidación, rendimiento, escalabilidad, disponibilidad y aislamiento. Una *Máquina Virtual* (VM: Virtual Machine) es un entorno de ejecución aislado creado por el software encargado de gestionar todos los recursos del computador al que se lo conoce como *Hipervisor* o *Monitor de Máquina Virtual* (VMM: Virtual Machine Monitor). Generalmente, el término VM se relaciona a las tecnologías de Virtualización de Hardware y de Paravirtualización, aunque también es utilizado en la virtualización de procesos tales como la *Máquina Virtual Java* (JVM: Java Virtual Machine). Los entornos de ejecución aislados creados a nivel de OS se denominan *Contenedores, Prisiones* o *Zonas* (según el OS). A los efectos

de utilizar una única denominación en el resto del resumen, a estos entornos o dominios de ejecución se los referirá como *Contenedores*.

Sin considerar cuál es la denominación del *entorno de ejecución* creado por la virtualización según su tipo, queda claro que éste resulta del particionado de los recursos de un computador. Como consecuencia de esto, *el poder de cómputo y la utilización de recursos de esos entornos estarán acotados por el computador que lo contiene*. Al menos dos preguntas surgen al considerar esta limitación:

1. ***¿Cómo se pueden expandir el poder de cómputo y la utilización de recursos de una VM o Contenedor a varios computadores?***

   Se pretende disponer de un entorno de ejecución aislado equivalente a una VM o Contenedor, pero que su alcance no se limite a un único computador.

2. ***¿Cómo lograr mayores niveles de rendimiento y escalabilidad de las aplicaciones que se ejecutan en la Nube?***

   Las aplicaciones que ejecutan en la Nube demandan cada vez más potencia de cómputo y recursos (memoria, almacenamiento, red, etc.). El rendimiento y la escalabilidad requerida son imposibles de lograr con un único computador. Además, dada la criticidad de las aplicaciones, existen otros requerimientos que deben cumplirse de tal modo que se les permita a éstas expandirse o contraerse según la demanda (elasticidad), tolerar fallos, ahorrar energía y facilitar su despliegue y gestión.

Los diseñadores de aplicaciones para la Nube utilizan metodologías de desarrollo de software basada en la Arquitectura de Micro-Servicios (MSA: Micro Service Architecture) [2] o Arquitectura Orientada al Servicio (SOA: Service Oriented Architecture) [3]. Ambas arquitecturas proponen dividir las aplicaciones en múltiples servicios desplegados en múltiples VMs o Contenedores. Para realizar la configuración y el despliegue de los diferentes componentes se utilizan gestores de Contenedores tales como Docker [4], Mesos [5] o Kubernetes [6]. En el caso de Mesos, éste se anuncia como un DC/OS (Data Center Operating System), pero ésto resulta confuso ya que en realidad no es un OS, al menos en lo que a la definición tradicional de un OS refiere. Mesos es el software que permite gestionar todos los recursos de un cluster, tratando de integrar la aplicación a esos recursos.

Por otro lado, cuando dos procesos componentes de la aplicación deben comunicarse, si los mismos se encuentran ejecutando en la misma VM o Contenedor (bajo el control del mismo OS) se pueden comunicar utilizando pipes o tuberías, memoria compartida, colas de mensajes, señales, semáforos, etc. Pero, cuando se comunican con procesos que se encuentran en otras VMs o Contenedores deben utilizar protocolos de red, sean éstos de alto nivel como HTTP, de nivel medio como RPC, o de bajo nivel como sockets TCP o UDP.

Si bien existen numerosas herramientas y una multiplicidad de software para desarrollar aplicaciones para la Nube, no conforman una solución transparente, consistente y homogénea como en un sistema centralizado. La complejidad recae entonces en el usuario, programador o administrador, el cual debe hacerse cargo de aspectos relativos al sistema y a los recursos distribuidos [7]. Esto se debe, a que todo el conjunto de recursos dispersos (de cómputo, de almacenamiento y de red) presentan una visión no homogénea que el usuario debe integrar, gestionar y mantener. Esto significa que, desde el punto de vista netamente práctico, se requiere realizar tareas tales como: asignar direcciones IP, puertos TCP/UDP, configurar reglas en los firewalls, establecer servidores de autenticación, mantener versiones de OS, mantener actualizadas bibliotecas y paquetes de software, y mantener archivos de configuración sincronizados entre VMs/Contenedores.

Cuando se utiliza IaaS para desplegar las aplicaciones de esta manera, no se puede asegurar que los Contenedores o las VMs contratadas se localicen en nodos diferentes con lo cual, muchos

supuestos realizados para distribuir cargas o para tolerar fallos a través de la redundancia dejan de tener sustento.

La tesis propone un modelo de arquitectura que se basa en este enfoque distribuyendo procesos, servicios y recursos para proveer un entorno de ejecución aislado basado en múltiples *Contenedores Distribuidos* (DC: Distributed Containers) y en la factorización de los OSs. El resultado es un Sistema de Virtualización Distribuido que combina tecnologías de OS y de virtualización de tal forma de que las aplicaciones puedan acceder a los recursos de su OS sin importar en el host donde se éstas se ejecutan. Esta característica simplifica el desarrollo de aplicaciones (costo y tiempo).

MOTIVACIÓN

Conforme una VM o Contenedor resulta de una partición de recursos de un computador, su potencia de cómputos y disponibilidad de recursos están limitadas al computador que lo contiene (las partes no pueden ser mayores que el todo). Esto nos permite realizar la siguiente afirmación:

En los sistemas de virtualización de la actualidad, los recursos y el poder de cómputo de una VM o de un Contenedor se encuentran limitados al computador donde residen.

En base a esta limitante, el modelo de arquitectura propuesto por la tesis permite que diferentes procesos que se ejecutan en diferentes nodos de un cluster pero que pertenecen al mismo entorno de ejecución (DC) utilicen recursos tal como si ejecutarían en un sistema centralizado (Imagen de Sistema Unico – SSI) constituyendo un nuevo modelo de arquitectura de virtualización.

Tesis: Se propone un modelo de Arquitectura de un Sistema de Virtualización Distribuido (DVS: Distributed Virtualization System), el cual permite construir dominios de ejecución denominados Contenedores Distribuidos (DC: Distributed Containers) que pueden extenderse más allá de los límites de una máquina física.

El objetivo es que las aplicaciones desplegadas en DCs dispongan de mayores cantidades de recursos y mayor poder de cómputo para poder brindar niveles de rendimiento, disponibilidad, escalabilidad y elasticidad más elevados. Los recursos (tanto físicos como abstractos) y procesos que constituyen un DC, pueden encontrase dispersos (y eventualmente replicados) en diferentes nodos de un cluster de virtualización.

Un DVS permite ocultar tanto a usuarios como a las aplicaciones cuáles son los nodos que componen un DC, cuántos de esos componentes se encuentran replicados, e incluso cuáles componentes se activan/desactivan cuando se producen fallos. Los usuarios y las aplicaciones de cada uno de los OSs que ejecutan en el entorno de un DC perciben un SSI.

El modelo propuesto conserva las características tan apreciadas de las tecnologías de virtualización actuales tales como confinamiento, consolidación y seguridad, y los beneficios de los DOSs tipo SSI, tales como la transparencia, mejor rendimiento, mayor elasticidad, disponibilidad y escalabilidad.

Tanto las aplicaciones legadas como las aplicaciones nativas para la Nube se benefician con sus características. En general, la migración de aplicaciones legadas que se ejecutan en servidores propios (on-premise) hacia aplicaciones desarrolladas para ejecutar en la Nube requiere de una importante inversión de dinero y tiempo. Si en la Nube estuviese disponible una interface estándar ofrecida por un VOS (tal como POSIX), pero sobre una infraestructura con capacidades extendidas a varios nodos de un cluster, las tareas de migración se simplificarían reduciendo los costos y tiempos.

Las aplicaciones desarrolladas bajo el modelo de MSA [2] o SOA [3] también se verían beneficiadas con un DVS. Un DVS es inherentemente distribuido, por lo que dispone de herramientas de comunicación que facilitan el intercambio de datos y la sincronización entre

servicios o micro-servicios. Un DVS presenta una vista única integral del conjunto de componentes de la aplicación que facilitan su despliegue, mantenimiento y operación.

Desde los inicios del proyecto, se comenzó con el desarrollo de un prototipo de DVS que permitiera validar diferentes aspectos considerados en el modelo, detectar errores de diseño y proponer mejoras. Se describen en la tesis los detalles de implementación de los diferentes componentes del prototipo y se presentan resultados de las mediciones previas a las optimizaciones. El prototipo demuestra la factibilidad de implementar un DVS, mediante el desarrollo módulos de kernel, bibliotecas, programas de pruebas, etc. para Linux. También se han convertido dos OSs a Virtual OSs (VOS) y se han utilizado herramientas y facilidades ya existentes en Linux.

CONTRIBUCIÓN

La contribución esperada de la tesis es un modelo de arquitectura de un DVS que permite conformar entornos de ejecución aislados (DCs) que aprovechan de la agregación de recursos computacionales de los nodos de un cluster. Como se pueden adoptar múltiples estrategias, enfoques, mecanismos y variantes para el modelo del DVS, la tesis se limita en su alcance a desarrollar aquellos componentes considerados centrales en importancia. Otras características y cualidades potenciales del modelo se proponen para futuros trabajos de investigación.

Además del modelo de arquitectura de DVS, en la tesis se realizan otras contribuciones a saber:

- ✓Un mecanismo de IPC que es independiente de la localización de los procesos (se encuentren los mismos en el mismo o en diferentes nodos), que soporta en forma transparente la migración de procesos, la distribución de cargas y la utilización de mecanismos de replicación de tipo Primario-Respaldo o tipo Máquina de Estado.
- ✓Un algoritmo distribuido tolerante a fallos de asignación de recursos (utilizado en el VOS distribuido del prototipo), que permite asignar recursos y mantener la consistencia y gran parte de la disponibilidad tanto en fallos de tipo de Detención Total o de Partición de Red.
- ✓Un VOS distribuido compuesto de varios procesos servidores dispersos en varios nodos del cluster, algunos de ellos replicados.
- ✓Un VOS tipo unikernel que contiene un servidor web, un sistema de archivos y una pila de protocolos TCP/IP del cual se pueden separar varios componentes hacia otros nodos, tales como el servidor de archivos o el servidor de almacenamiento en disco.

Más allá de prototipo construido, para el lector debe quedar claro que el modelo propuesto es de carácter teórico y genérico, el cual puede extenderse, modificarse o simplemente ser inspirador para otros tipos de enfoques.

ALCANCE DE LA TESIS

El alcance de la tesis abarca la descripción del modelo de arquitectura de un DVS que se propone como una nueva tecnología de virtualización, la descripción de sus componentes, el análisis de los beneficios que se obtienen con este sistema, la descripción del prototipo de DVS y una evaluación de rendimiento del mismo.

Aspectos relacionados a su integración con arquitecturas de gestión de servicios en la Nube quedan más allá del alcance. Estos y otros muchos tópicos relacionados quedan abiertos para futuros trabajos de investigación.

TECNOLOGIAS RELACIONADAS

El modelo de arquitectura de un DVS surge a partir de integrar múltiples tecnologías de virtualización, de OSs y de sistemas distribuidos en general. Todas ellas fueron fuentes de inspiración para construir el modelo propuesto.

1) *Sistemas Operativos Multiservidor o de Microkernel* [8]: A diferencia de los OS monolíticos, los OS multiservidor basados en microkernel descomponen al sistema en múltiples capas y múltiples procesos aislados que ejecutan en modo usuario. Estos procesos se comunican entre sí mediante mecanismos de IPC que provee el microkernel. El microkernel brinda servicios básicos a las capas superiores y es el único componente que ejecuta en modo kernel o modo supervisor.

2) *Sistemas Operativos Distribuidos con Imagen de Sistema Unica [9]*: Estos OSs distribuyen la ejecución de procesos en diferentes nodos de un cluster ofreciendo al usuario una imagen única de sistema. Algunos de estos sistemas permiten la migración de procesos para distribución y/o balanceo de cargas.

3) *Virtualización de Sistema Operativo - Sistemas Operativos Virtuales [10]*: Un OS ofrece servicios y abstracciones a las aplicaciones a través de su interface superior, y se comunica con el hardware de la plataforma a través de su interface inferior. Al igual que un OS, un VOS ofrece servicios y abstracciones a las aplicaciones a través de su interface superior, pero utiliza los servicios de un OS subyacente a través de su interface inferior. El funcionamiento es similar al de un OS ejecutando sobre un sistema paravirtualizado, excepto que utiliza Llamadas al Sistema (syscalls) en vez de Llamadas al Hipervisor (hypercalls).

4) *Virtualización basada en Sistema Operativo con Contenedores [11]:* El kernel del OS-Host particiona sus recursos en instancias aisladas en espacio de usuario. El sistema de virtualización construye un Contenedor que confina a un conjunto de aplicaciones y a su entorno de ejecución. Esta tecnología se basa en la virtualización de Llamadas al Sistema. Si bien todas estas instancias se ejecutan en sus propios entornos, todas comparten el mismo OS.

5) *Sistema de Comunicaciones Grupales (GCS) [12]:* Como el DVS es un sistema distribuido, se requiere de mecanismos de comunicación para sus diferentes componentes, y una alternativa es utilizar un GCS. Dentro de las posibilidades que ofrecen los sistemas distribuidos se encuentra la replicación; al utilizar GCS ésta debe ofrecer diferentes opciones de entrega de mensajes tales como FIFO, Causal, Orden Total, etc. Por otro lado, como los sistemas distribuidos que ejecutan en un cluster, la falla de uno de los nodos del mismo puede provocar una falla completa del sistema. Para reducir esta posibilidad, el sistema de GCS debe ofrecer a las aplicaciones mecanismos de tolerancia a fallos.

La descomposición (factorización) en múltiples tareas y servidores comunicándose mediante mecanismos de IPC como proponen los OS de microkernel brinda el aislamiento requerido por un sistema de virtualización y permite la distribución de los mismos en diferentes computadores siempre y cuando se puedan comunicar mediante IPC y GCS.

Actualmente los servicios de infraestructura proveen recursos tales como CPUs, memoria, disco y red, pero el usuario carece de una visión integrada tal como la que ofrece un OS multikernel debiendo gestionar explícitamente sus recursos en la Nube. El modelo de arquitectura de DVS propone una gestión integrada de recursos y transparencia para las aplicaciones. Estas características, simplifican la programación, el despliegue y mantenimiento de aplicaciones en la Nube.

MODELO DE ARQUITECTURA PROPUESTO

TOPOLOGÍA DE UN CLUSTER DVS

El modelo de arquitectura de DVS cumple con los requerimientos de los proveedores de Servicios en la Nube tales como alta disponibilidad, replicación, elasticidad, balanceo de carga, gestión de recursos y migración de procesos entre otros. Su objetivo principal es aumentar el rendimiento asignando recursos de un subconjunto de nodos a cada instancia de un VOS (agregación de recursos), pero permitiendo que múltiples VOS compartan los mismos nodos

(partición de recursos). Cada VOS se ejecuta en un entorno aislado denominado Contenedor Distribuido (DC), el cual puede abarcar desde uno a todos los nodos del cluster DVS. Un DC es un entorno de ejecución aislado y distribuido que ofrece recursos computacionales, virtuales y abstractos a un conjunto de procesos. En la Fig. 1, se presenta un ejemplo de topología de un cluster DVS que cuenta con cinco nodos unidos por una red de alta velocidad.

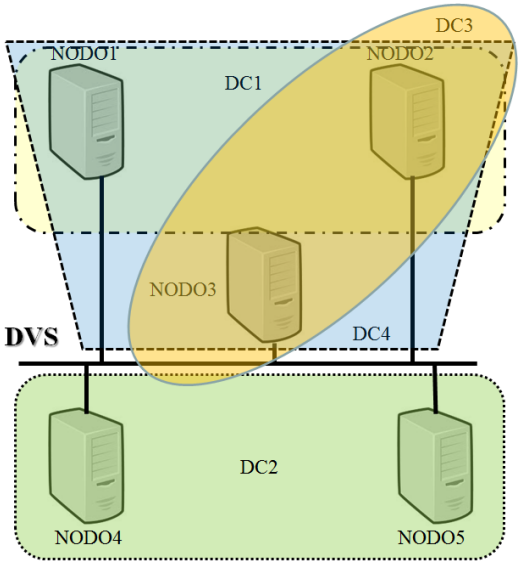


Figura 1.Ejemplo de Topología de un cluster DVS.

MODELO DE ARQUITECTURA DE DVS

El modelo de arquitectura de DVS está conformado por tres capas (Fig. 2). En su capa inferior se encuentran el hardware, con su OS-host que tiene embebido un módulo que conforma uno de los componentes del DVK.

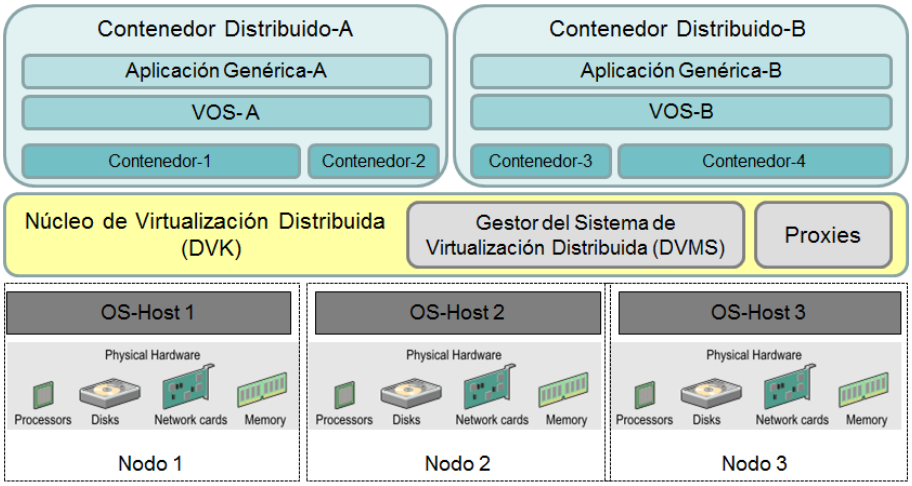


Figura 2.Modelo de Arquitectura del ***DVS***.

La capa intermedia la conforman el DVK, el Sistema de Gestión de Virtualización Distribuida (DVMS: Distributed Virtualization Management System) y los proxies. El DVK está compuesto por una serie de servicios que se ofrecen a la capa superior a través de sus APIs (Fig. 3), y los cuales utilizan servicios de la capa inferior a través del módulo embebido. El DVMS dispone de las APIs, comandos, bibliotecas, interfaces gráficas o interfaces web para gestionar el DVS y todos sus componentes. Cada servicio del DVK que se ofrece a las capas superiores (GCS, exclusión mutua distribuida, consenso distribuido, etc.) debe tener en consideración que los

procesos que involucran deben pertenecer al mismo DC, tal como lo hace el mecanismo de IPC. Es función del DVK controlar el aislamiento entre DCs para todos estos servicios.

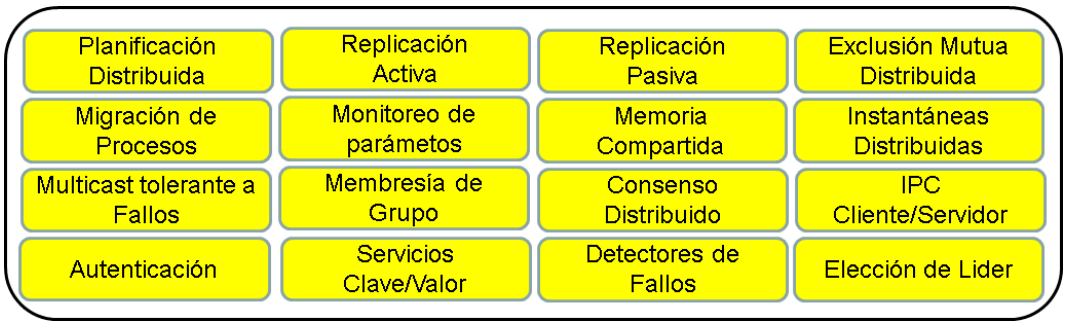


Figura 3. Servicios del Distributed Virtualization Kernel (*DVK*).

Los proxies son los procesos que transfieren mensajes y datos entre los nodos del cluster del DVS. Finalmente, la capa superior la conforman los DCs, los Contenedores, los VOS y las aplicaciones.

Núcleo de Virtualización Distribuida (DVK)

El DVK es la capa de software principal que integra los recursos del cluster, gestiona y limita los recursos asignados a cada DC. Ofrece a las capas superiores las interfaces necesarias para utilizar los protocolos y servicios de bajo nivel (Fig. 3), los cuales pueden ser utilizados por los VOS, tales como Planificación Distribuida, Replicación Activa/Pasiva, Exclusión Mutua Distribuida, mecanismos de migración de procesos, monitoreo de parámetros de rendimiento, Memoria Compartida Distribuida, Instantáneas Distribuidas, Multicast tolerante a Fallos, Membresía de Grupos, Consenso Distribuido, IPC Cliente/Servidor, servicios de clave/valor, autenticación, detección de fallos, elección de Lider, etc.

Representantes de Comunicaciones de Nodos Remotos (Proxies)

El cluster donde se ejecutará el DVS se define básicamente utilizando nodos y proxies. Los proxies de comunicaciones son los procesos responsables de transferir mensajes y bloques de datos entre un nodo Local y uno o más nodos remotos. Lo existencia de proxies se oculta al resto de los componentes del DVS tales como DCs, Contenedores y procesos registrados.

Contenedores

Los Contenedores son entornos de ejecución aislados que brinda el OS. Entre sus características significativas para su utilización en el DVS se destaca el aislamiento y el espacio de nombres (*namespace*) propio. El aislamiento de rendimiento se logra limitando ciertos recursos del OS asignados a cada Contenedor tales como tiempo de CPU, buffers, caches, reserva de ancho de banda de la red, E/S de dispositivos, memoria, etc.

Contenedores Distribuidos

Un DC también es un conjunto de Contenedores dispersos en diferentes nodos de un cluster, pero que permite comunicar los procesos de un Contenedor con los procesos de otro Contenedor en diferente nodo utilizando un IPC transparente a la localización. El mecanismo de IPC solo admite la comunicación entre procesos que se ejecutan en un mismo DC, estén estos localizados en el mismo o en diferentes nodos del cluster de virtualización.

Sistema Operativo Virtual

Se necesita un VOS para que las aplicaciones no pueden realizar Llamadas al Sistema al OS-host en forma directa, y para que puedan ejecutarse diferentes VOS sobre el mismo OS-host. Se admiten

diferentes tipos de VOS tales como monolíticos, Unikernels o Distribuidos los cuales pueden ejecutar concurrentemente en diferentes DCs.

PROTOTIPO

Prácticamente, en forma simultánea al diseño del modelo de arquitectura de DVS, se comenzó con el desarrollo de un prototipo con el fin de confirmar la factibilidad de la propuesta, detectar errores de diseño y proponer mejoras. Los componentes que lo constituyen son:

- *DVK*: En él se implementan, por un lado, los mecanismos de IPC (y sus APIs) extendidos para transferir mensajes y datos entre nodos de un cluster, para soportar migración y esquemas de replicación. Por otro lado, se implementaron los mecanismos de gestión del DVS tales como iniciar y detener el DVS, configurar, iniciar y detener DCs, registrar y de-registrar nodos y procesos proxies, registrar y de-registrar procesos ordinarios dentro de un DC, obtener información de estado y estadística de cada componente del DVS.
- *GCS*: Por cuestiones prácticas, en el DVK no se incorporan comunicaciones grupales, aquellas aplicaciones que las requerían utilizan Spread Toolkit [13].
- *VOS-multiservidor*: Como se disponía de la base del desarrollo de MoL (Minix Over Linux) se convirtieron sus componentes en los componentes de un VOS-multikernel,
- *VOS-unikernel:* Como base para el desarrollo de un VOS-unikernel se tomó el software LwIP [14], el cual dispone de una pila de protocolo TCP/IP completa, un servidor web básico con página estática en RAM, e interfaces de LAN virtual
- *Gestión por Webmin*: Si bien el prototipo puede ser gestionado mediante Interface de Línea de Comandos (CLI: Command Line Interface), o utilizando las APIs del DVMS que forman parte del DVK, se desarrolló un módulo Webmin que permite gestionar el DVS.

CONCLUSIONES

Los proveedores de IaaS siempre requieren de altos niveles de rendimiento, escalabilidad y disponibilidad para brindar sus servicios de virtualización. Estos requerimientos se pueden satisfacer con una tecnología de virtualización distribuida. El modelo de arquitectura de DVS propuesto combina e integra las tecnologías de virtualización y de los Sistemas Distribuidos para proporcionar los beneficios de ambos mundos, lo que lo hace adecuado para ofrecer Servicios en la Nube con nivel de proveedor. La arquitectura se basa en poner a disposición de los VOS y aplicaciones mecanismos de IPC, GCS y servicios autónomos, distribuidos y débilmente acoplados otorgándoles flexibilidad para los desarrollos de nuevos servicios y aplicaciones, pero manteniendo las características de aislamiento requeridas.

Un DVS es que facilita la migración/portación de aplicaciones heredadas desde servidores locales (on-premise) a la Nube. Con un VOS que suministre las API POSIX y RPC permitiría la reutilización de gran parte del código de esas aplicaciones, mejorando los costos y tiempos de implementación al reducir el esfuerzo de codificación. Por otro lado, también se adapta perfectamente a las aplicaciones basadas en el MSA o SOA, éstas pueden ejecutar un conjunto de microservicios en varios nodos del clúster mediante el uso de IPC/RPC propio del DVS para las comunicaciones.

Con un DVS se da respuesta a las dos preguntas planteadas en la Introducción: 1) Los límites de un entorno aislado para la ejecución de aplicaciones (DC) pueden expandirse a todos los nodos del clúster (agregación) mejorando sus características de rendimiento, escalabilidad y disponibilidad. Además, los nodos de un cluster pueden compartirse entre varios DCs (partición), mejorando así la utilización de la infraestructura de hardware. 2) Un DVS soporta un VOS distribuido ejecutando dentro de un DC el cual le aporta una visión integrada de las aplicaciones equivalente a un sistema centralizado. Esta visión integrada facilita la gestión reduciendo los tiempos y los costos de implementación, operación y mantenimiento de aplicaciones en la Nube.

El modelo de arquitectura propuesto abre la posibilidad de investigar y desarrollar varios aspectos relacionados tanto al DVS propiamente dicho, como así también acerca de algunos de sus componentes, tales como los VOSs.

TRABAJOS FUTUROS

Para los próximos años se planea trabajar en los siguientes proyectos de investigación y desarrollo (los cuales están actualmente en curso bajo el PID 5162 – UTN - FRSF) que pueden producir un gran impacto en la adopción del modelo de DVS:

1) Incorporar estándares de gestión tales como Openstack[15],

2) Incorporar a UML (User Mode Linux) como VOS,

3) Incorporar un VOS Unikernel generado con Rumpkernel [16] con soporte POSIX.

4) Desarrollar un Router Virtual Distribuido basado en Contrail vRouter [17] que utilice Virtual Network Interface Cards (vNICs) localizadas en diferentes nodos del cluster DVS.

El éxito en estos proyectos aumentaría la confianza para la adopción del modelo propuesto.

REFERENCIAS

[1] P. M. Mell, T. Grance, "*The NIST Definition of Cloud Computing*", Technical Report, NIST, SP 800-145, 2011.

[2] C. Pautasso, O. Zimmermann, M. Amundsen, J. Lewis, and N. Josuttis, "*Microservices in Practice, Part 1: Reality Check and Service Design*", IEEE Softw. 34, pp 91-98, Jan. 2017, 2017.

[3] N. Bieberstein et al., "*Service-Oriented Architecture Compass*", Pearson, ISBN 0-13-187002-5, 2006.

[4] J. Turnbull, "*The Docker Book*", 2014, Available online at: https://www.dockerbook.com/, accedido en Enero 2018.

[5] B. Hindman, A. Konwinski, M. Zaharia, A. Ghodsi, A. D. Joseph, R. Katz, S. Shenker, and Ion Stoica, "*Mesos: a platform for fine-grained resource sharing in the data center*", Proc. of the 8th USENIX conference on Networked systems design and implementation (NSDI'11), USENIX Association, Berkeley, CA, USA, pp. 295-308, 2011.

[6] N. Poulton, "*The Kubernetes Book*", ISBN-13: 978-1521823637, ISBN-10: 1521823634 ,2017.

[7] D. Wentzlaff, C. Gruenwald III, N. Beckmann, K. Modzelewski, A. Belay, L. Youseff, J. Miller, A. Agarwal, "*An operating system for multicore and clouds: mechanisms and implementation*", in Proceedings of the 1st ACM symposium on Cloud computing (SoCC '10). ACM, New York, NY, USA, 3-14, 2010.

[8] A. Tanenbaum, et al, "*Minix 3: Status Report and Current Research*", login: The USENIX Magazine, 2010.

[9] OpenSSI (Single System Image) Clusters for Linux, http://www.openssi.org/cgi-bin/view?page=openssi.html , accedido Enero 2018.

[10] D. Hall, D. Scherrer, J. Sventek, "*A Virtual Operating System*", Journal Communication of the ACM, 1980.

[11] G. Banga, P. Druschel, J. C. Mogul, "*Resource Containers: A New Facility for Resource Management in Server Systems*", in Operating Systems Design and Implementation, 1999.

[12] K. P. Birman, "*The process group approach to reliable distributed computing*", Commun. ACM 36, Dec. 1993, pp. 37-53, doi:http://dx.doi.org/10.1145/163298.163303.

[13] The Spread Toolkit. http://www.spread.org , accedido Enero 2018.

[14] LwIP, http://savannah.nongnu.org/projects/lwip/ , accedido Enero 2018.

[15] Openstack, https://www.openstack.org/ , accedido Enero 2018.

[16] Rumpkernel, http://rumpkernel.org/ , accedido Enero 2018.

[17] vRouter Contrail, https://github.com/Juniper/contrail-vrouter, accedido Enero 2018.