



ISSN: 1646-9895

Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

O u t u b r o 1 7 • O c t o b e r 1 7



©AISTI 2017 <http://www.aisti.eu>

Nº 24

Edição / Edición

Nº 24, 10/2017

Tiragem / Tirage: 1000

Preço por número / Precio por número: 17,5€

Subscrição anual / Suscripción anual: 30€ (2 números)

ISSN: 1646-9895

Depósito legal:

Indexação / Indexación

Academic Journals Database, CiteFactor, Dialnet, DOAJ, DOI, EBSCO, GALE, IndexCopernicus, Index of Information Systems Journals, ISI Web of Knowledge, Latindex, ProQuest, QUALIS, SciELO, SCImago, Scopus, SIS, Ulrich's.

Propriedade e Publicação / Propriedad y Publicación

AISTI – Associação Ibérica de Sistemas e Tecnologias de Informação

Rua Quinta do Roseiral 76, 4435-209 Rio Tinto, Portugal

E-mail: aistic@gmail.com

Web: <http://www.aisti.eu/>

Ficha Técnica

Director

Álvaro Rocha, Universidade de Coimbra

Coordenadores da Edição / Coordinadores de la Edición

Gloria Piedad Gasca Hurtado, Universidad de Medellín

Conselho Editorial / Consejo Editorial

Carlos Ferrás Sexto, Universidad de Santiago de Compostela

Gonçalo Paiva Dias, Universidade de Aveiro

Jose Antonio Calvo-Manzano Villalón, Universidad Politécnica de Madrid

Luís Paulo Reis, Universidade do Minho

Manuel Pérez Cota, Universidad de Vigo

Ramiro Gonçalves, Universidade de Trás-os-Montes e Alto Douro

Conselho Científico / Consejo Científico

Adolfo Lozano-Tello, Universidad de Extremadura, ES

Adriano Pasqualotti, Universidade de Passo Fundo, BR

Alberto Fernández, Universidad Rey Juan Carlos, ES

Alberto Bugarín, Universidad de Santiago de Compostela, ES

Alejandro Medina, Universidad Politécnica de Chiapas, MX

Alejandro Rodríguez González, Universidad Politécnica de Madrid, ES

Alejandro Peña, Escuela de Ingeniería de Antioquia, CO

Alexandre L'Erario, Universidade Tecnológica Federal do Paraná, BR

Alma María Gómez-Rodríguez, Universidad de Vigo, ES

Álvaro E. Prieto, Universidad de Extremadura, ES

Ana Azevedo, Instituto Politécnico do Porto, PT

Ana Cristina Ramada Paiva, FEUP, Universidade do Porto, PT

Ana Isabel Veloso, Universidade de Aveiro, PT

Ana Maria Correia, ISEGI, Universidade Nova de Lisboa, PT

Ana Paula Afonso, Instituto Politécnico do Porto, PT

Anabela Mesquita, Instituto Politécnico do Porto, PT

Angelica Caro, Universidad del Bío-Bío, CL

Ania Cravero, Universidad de La Frontera, CL

Antoni Lluís Mesquida Calafat, Universitat de les Illes Balears, ES
Antonia Mas Pichaco, Universitat de les Illes Balears, ES
António Coelho, FEUP, Universidade do Porto, PT
António Godinho, ISLA-Gaia, PT
Antonio Jesus Garcia Loureiro, Universidad de Santiago de Compostela, ES
António Pereira, Instituto Politécnico de Leiria, PT
Armando Mendes, Universidade dos Açores, PT
Arnaldo Martins, Universidade de Aveiro, PT
Arturo J. Méndez, Universidad de Vigo, ES
Baltasar García Pérez-Schofield, Universidad de Vigo, ES
Benjamim Fonseca, Universidade de Trás-os-Montes e Alto Douro, PT
Bráulio Alturas, ISCTE - Instituto Universitário de Lisboa, PT
Brenda L. Flores-Rios, Universidad Autónoma de Baja California, MX
Brígida Mónica Faria, ESTSP, Instituto Politécnico do Porto, PT
Carlos Costa, Universidade de Aveiro, PT
Carlos Rabadão, Instituto Politécnico de Leiria, PT
Carlos Carreto, Instituto Politécnico da Guarda, PT
Carlos Morais, Instituto Politécnico de Bragança, PT
Carlos Vaz de Carvalho, Instituto Politécnico do Porto, PT
Carmen Galvez, Universidad de Granada, ES
César Gonçalves, Universidade do Algarve, PT
Ciro Martins, Universidade de Aveiro, PT
Cristina Alcaraz, Universidad de Málaga, ES
Daniel Castro Silva, Universidade de Coimbra, PT
Daniel Polónia, Universidade de Aveiro, PT
Daniel Riesco, Universidad Nacional de San Luis, AR
David Fonseca, Universitat Ramon Llull, ES
Dora Simões, Universidade de Aveiro, PT
Emiliano Reynares, CIDISI - UTN FRSF - CONICET, AR
Eusébio Ferreira da Costa, Escola Superior de Tecnologias de Fafe, PT
Feliz Gouveia, Universidade Fernando Pessoa, PT
Fernando Bandeira, Universidade Fernando Pessoa, PT
Fernando Diaz, Universidad de Valladolid, ES
Fernando Moreira, Universidade Portucalense, PT

Francisco Restivo, Universidade Católica Portuguesa, PT
Gerardo Gonzalez Filgueira, Universidad da Coruña, ES
Gerardo Rodriguez, Universidad de Salamanca, ES
Germano Montejano, Universidad Nacional de San Luis, AR
Guilhermina Lobato Miranda, Universidade de Lisboa, PT
Hélder Zagalo, Universidade de Aveiro, PT
Hélia Guerra, Universidade dos Açores, PT
Henrique Gil, Instituto Politécnico de Castelo Branco, PT
Henrique Santos, Universidade do Minho, PT
Higino Ramos, Universidad de Salamanca, ES
Hugo Paredes, Universidade de Trás-os-Montes e Alto Douro, PT
Isabel Pedrosa, Instituto Politécnico de Coimbra, PT
Isaura Ribeiro, Universidade dos Açores, PT
Isidro Calvo, Universidad del País Vasco (UPV/EHU), ES
Ismael Etxeberria-Agiriano, Universidad del País Vasco (UPV/EHU), ES
Ivan Garcia, Universidad Tecnológica de la Mixteca, MX
Javier Garcia Tobio, CESGA-Centro de Supercomputacion de Galicia, ES
Jezreel Mejia, Centro de Investigación en Matemática (CIMAT), MX
João Pascual Faria, FEUP, Universidade do Porto, PT
João Paulo Costa, Universidade de Coimbra, PT
João Tavares, FEUP, Universidade do Porto, PT
Joaquim José Gonçalves, Instituto Politécnico do Cávado e do Ave, PT
Joaquim Madeira, Universidade de Aveiro, PT
Joaquim Reis, ISCTE - Instituto Universitário de Lisboa, PT
Jörg Thomaschewski, University of Applied Sciences OOW - Emden, DE
José Augusto Fabri, Universidade Tecnológica Federal do Paraná, BR
José Braga de Vasconcelos, Universidade Atlântica, PT
José Felipe Cocón Juárez, Universidad Autónoma del Carmen, MX
Jose J. Pazos-Arias, Universidad de Vigo, ES
José Luís Silva, Universidade da Madeira, PT
José Paulo Lousado, Instituto Politécnico de Viseu, PT
José Luis Pestrana Brincones, Universidad de Málaga
José Luís Reis, ISMAI - Instituto Superior da Maia, PT
Jose M Molina, Universidad Carlos III de Madrid, ES

Jose Maria Zavala Perez, Eticas Research & Consulting, ES
José Martins, Universidade de Trás-os-Montes e Alto Douro, PT
Jose R. R. Viqueira, Universidade de Santiago de Compostela, ES
José Silvestre Silva, Academia Militar, PT
Josep M. Marco-Simó, Universitat Oberta de Catalunya, ES
Juan D'Amato, PLADEMA-UNCPBA-CONICET, AR
Juan M. Santos Gago, Universidad de Vigo, ES
Juan Manuel Fernández-Luna, Universidad de Granada, ES
Juan-Manuel Lopez-Zafra, Universidad Complutense de Madrid, ES
Leonardo Bermon, Universidad Nacional de Colombia, CO
Leila Weitzel, Universidade Federal Fluminense, BR
Lilia Muñoz, Universidad Tecnológica de Panamá, PA
Luis Alvarez Sabucedo, Universidad de Vigo, ES
Luís Correia, Universidade de Lisboa, PT
Luis de Campos, Universidad de Granada, ES
Luis Enrique, Sicaman Nuevas Tecnologías S.L., ES
Luis Fernandez-Sanz, Universidad de Alcalá, ES
Luis Vilán-Crespo, Universidad de Vigo, ES
Luisa María Romero-Moreno, Universidad de Sevilla, ES
Luisa Miranda, Instituto Politécnico de Bragança, PT
Lus Sussy Bayona Ore, Universidad Nacional Mayor de San Marcos, PE
Magdalena Arcilla Cobián, Universidade Nacional de Educação a Distancia, ES
Marcelo Mendonça Teixeira, Universidade Federal Rural de Pernambuco, BR
Marco Painho, ISEGI, Universidade Nova de Lisboa, PT
Maria Hallo, Escuela Politécnica Nacional, EC
María J. Lado, Universidad de Vigo, ES
Maria João Castro, Instituto Politécnico do Porto, PT
Maria João Ferreira, Universidade Portucalense, PT
Maria João Gomes, Universidade do Minho, PT
Maria José Angélico, Instituto Politécnico do Porto, PT
Maria José Marcelino, Universidade de Coimbra, PT
Maria José Sousa, Universidade Europeia, PT
Marisol B. Correia, Universidade do Algarve, PT
Maristela Holanda, Universidade de Brasília, BR

Martín Llamas Nistal, Universidad de Vigo, ES
Mercedes Ruiz, Universidad de Cádiz, ES
Miguel A. Brito, Universidade do Minho, PT
Miguel Bugalho, Universidade Europeia, PT
Miguel Casquilho, IST, Universidade de Lisboa, PT
Mirna Ariadna Muñoz Mata, Centro de Investigación en Matemáticas (CIMAT), MX
Nelson Rocha, Universidade de Aveiro, PT
Nuno Lau, Universidade de Aveiro, PT
Nuno Ribeiro, Universidade Fernando Pessoa, PT
Orlando Belo, Universidade do Minho, PT
Oscar Mealha, Universidade de Aveiro, PT
Paula Peres, Instituto Politécnico do Porto
Paula Prata, Universidade da Beira Interior, PT
Paulo Martins, Universidade de Trás-os-Montes e Alto Douro, PT
Paulo Pinto, FCT, Universidade Nova de Lisboa, PT
Pedro Abreu, Universidade de Coimbra, PT
Pedro Miguel Moreira, Instituto Politécnico de Viana do Castelo, PT
Pedro Nogueira Ramos, ISCTE - Instituto Universitário de Lisboa, PT
Pedro Sánchez Palma, Universidad Politécnica de Cartagena, ES
Pedro Sanz Angulo, Universidad de Valladolid, ES
Pilar Mareca Lopez, Universidad Politécnica de Madrid, ES
Raul Laureano, ISCTE - Instituto Universitário de Lisboa, PT
Renata Spolon Lobato, UNESP - Universidade Estadual Paulista, BR
Reinaldo Bianchi, Centro Universitário da FEI, BR
Rita Santos, Universidade de Aveiro, PT
Roberto Rodríguez Echeverría, Universidad de extremadura, ES
Rogério Eduardo Garcia, Universidade Estadual Paulista, BR
Rubén González Crespo, Universidad Internacional de La Rioja, ES
Rui Cruz, IST, Universidade de Lisboa, PT
Rui José, Universidade do Minho, PT
Rui Pedro Marques, Universidade de Aveiro, PT
Santiago Gonzales Sánchez, Universidad Inca Garcilaso de la Vega, PE
Sergio Gálvez Rojas, Universidad de Málaga, ES
Sérgio Guerreiro, Universidade Lusófona de Humanidades e Tecnologias, PT

Silvia Fernandes, Universidade do Algarve, PT
Solange N Alves de Souza, Universidade de São Paulo, BR
Tomás San Feliu Gilabert, Universidad Politécnica de Madrid, ES
Valéria Farinazzo Martins, Universidade Presbiteriana Mackenzie, BR
Vitor Carvalho, Instituto Politécnico do Cávado e do Ave, PT
Vitor Santos, ISEGI, Universidade Nova de Lisboa, PT
Wagner Tanaka Botelho, Universidade Federal do ABC, BR
Xose A. Vila, Universidad de Vigo, ES

Ciencias de la Computación en el mundo real

Computer Science in the real world

Gloria Piedad Gasca-Hurtado ¹, Liliana Machuca-Villegas^{1,2}

gpgasca@udem.edu.co, liliana.machuca@correounivalle.edu.co

¹ Universidad de Medellín, Carrera 87 No. 30-65, 50026, Medellín, Colombia.

² Universidad del Valle, Calle 13 # 100-00, 760032, Cali, Valle del Cauca, Colombia.

DOI: 10.17013/risti.24.0

1. Introducción

Los artículos presentados en este número fueron seleccionados de los trabajos aceptados en el Décimo Seminario de Ciencias de la Computación (SICC) 2017, realizado en la Ciudad de Medellín, Antioquia, Colombia del 18 al 20 de octubre del 2017. Este seminario ha tenido una tasa de aceptación del 21% de los artículos enviados.

En este número de la Revista Ibérica de Sistemas y Tecnologías de la Información (RISTI), se abordan temas relacionados con las “Ciencias de la Computación”. Los trabajos de este número son las versiones originales presentadas por los autores en comunicación oral durante el congreso SICC 2017 y que fueron seleccionados después de realizarse un proceso de evaluación por pares doble ciego.

2. Ciencias de la Computación

Las Ciencias de la Computación es una disciplina que consiste en el estudio sistemático de la viabilidad, estructura, expresión y mecanización de los procesos algorítmicos que subyacen en la adquisición, representación, procesamiento, almacenamiento, comunicación y acceso a la información. Se puede encontrar que dicha información está codificada en bits y bytes en una memoria de computador o transcritos en genes y estructuras de proteínas en una célula humana (Ralston, Reilly, & Hemmendinger, 1976).

Hoy por hoy la computación es una herramienta que facilita la implementación de soluciones en el mundo real, ya sea de índole académica o científica, en áreas de la ciencia como la medicina, la biología, la informática o la computación.

Es bien sabido que las Ciencias de la Computación no se trata sólo de la construcción de computadores o el desarrollo de software (Denning, 2005). Se trata de una ciencia que busca poner al servicio del mundo las herramientas que existen en las diferentes disciplinas para la solución de problemas (Brookshear & Brookshear, 2002).

Aunque es posible que a través de la informática o de la construcción de software se aborde parte de un problema específico, es por medio de las Ciencias de la Computación que se consigue resolver problemas relativos a los procesos computacionales que pueden ser automatizados e implementados de manera eficiente (Newell, Perlis, & Simon, 1967).

Por consiguiente, los científicos en Ciencias de la Computación estudian la naturaleza misma de la informática para determinar qué problemas se pueden resolver. Es decir, las Ciencias de la Computación abarcan problemas más amplios que pueden estar inmersos en diferentes áreas de conocimiento, por ejemplo: es posible comparar varios algoritmos para determinar cuál proporciona la mejor solución en términos de rendimiento y/o precisión; o bien, diseñar lenguajes de programación para facilitar la especificación y expresión de algoritmos. También se puede decir que esta ciencia permite el diseño, evaluación y construcción de sistemas informáticos que pueden ejecutar eficientemente las especificaciones dadas (Brookshear & Brookshear, 2002), (Biermann, 1997).

En resumen, cuando se habla de Ciencias de la Computación se tienen en cuenta características como la proposición inherente de grandes soluciones, el uso del análisis matemático y el rigor lógico para evaluar dichas soluciones, el modelado de problemas complejos mediante el uso de abstracciones, el análisis e identificación de variantes del mismo problema en diferentes entornos y la proposición de soluciones eficientes conocidas a problemas en nuevas configuraciones (Brookshear & Brookshear, 2002), (Henderson, 2009).

En este sentido, los trabajos seleccionados para este número especial de la revista RISTI son un conjunto de reportes de investigación que presentan avances en diferentes áreas de conocimiento dentro de las que se encuentran las siguientes:

- La Inteligencia Artificial (IA) por ser una de las ramas de las Ciencias de la Computación aplicada. En los trabajos presentados se aborda la IA desde el punto de vista del uso de técnicas de procesamiento de imágenes y agrupamiento jerárquico.
- Internet de las Cosas como un vehículo para la solución de problemas en el área de la agricultura, específicamente para el análisis de variables climáticas en agricultura de precisión.
- La Innovación Abierta estudiada bajo un enfoque de co-creación de valor para que las organizaciones se beneficien del uso intensivo del conocimiento interno y externo, con el fin de potenciar y acelerar su innovación interna y expandir los mercados.
- La Ingeniería de Requisitos como una rama de la Ingeniería de Software que involucra, no solamente el conocimiento técnico sino los factores sociales y humanos asociados a procesos como la elicitación de requisitos.

Los artículos publicados en este número de RISTI han sido seleccionados porque abordan temáticas de interés en la actualidad y se enmarcan en soluciones del mundo real con un enfoque asociado a las Ciencias de la Computación.

3. Estructura

En el primer artículo se presenta el desarrollo de un nuevo medidor de concentraciones de cobre en línea, basado en el procesamiento de imágenes y la técnica de colorimetría. Los autores han utilizado algoritmos genéticos para mejorar la estimación del modelo de correlación planteado, obteniendo un producto que se diferencia de los existentes en su bajo costo y aportando significativamente a solucionar las dificultades de las plantas que carecen de estos medidores.

En el segundo artículo se utilizan técnicas de agrupamiento jerárquico para análisis de grupos puntuales. Es decir, en esta investigación se pretende analizar las diferencias en el espectrograma de los cantos de dos poblaciones de la misma especie de anuros, pero localizadas en diferentes lugares. Entre otros aspectos, el estudio presentado prueba que el monitoreo acústico puede detectar la influencia que el ambiente causa en las vocalizaciones de una misma especie.

En el tercer artículo se propone la aplicación de lógica difusa en la educación. Específicamente en este trabajo se establece un modelo difuso basado en 2-tuplas lingüísticas para evaluar competencias educativas. La principal motivación para realizar esta propuesta radica en la identificación de ciertas dificultades e imprecisiones para realizar la evaluación cuando el docente usa estructuras en las que se establecen relaciones, por la gran cantidad de información que se obtiene.

En el cuarto artículo se presenta una arquitectura para agricultura de precisión soportada en tecnologías abiertas para IoT y técnicas de minería de datos para la obtención, visualización y análisis en tiempo real de variables climáticas en un cultivo de prueba.

El quinto artículo define un protocolo para seleccionar una red social acorde a las necesidades específicas de un proceso de co-creación de valor, tomando como punto de partida que la co-creación hace alusión a la generación conjunta de valor a partir de la interacción entre agentes internos y externos a la organización. Se presenta una revisión de literatura para establecer un referente conceptual y lograr el análisis de la problemática. A partir de lo anterior, el trabajo logra un protocolo para seleccionar una red social y aportar en el proceso de co-creación de las organizaciones.

Finalmente, el sexto artículo hace énfasis en la solución de problemas de Ingeniería de Software estudiando no solamente el conocimiento técnico sino los factores sociales y humanos que afectan procesos de ingeniería de requisitos como la elicitación. La principal justificación de este trabajo radica en que aunque existen técnicas para elicitar los requerimientos de software en un contexto particular, no existe un conjunto amplio de estudios que abarquen aspectos sociales y humanos en el proceso de educación. Los hallazgos más relevantes del trabajo están fundamentados en la identificación de un conjunto de factores sociales y humanos que un especialista en requisitos debería poseer, adicionales a su disciplina, para realizar actividades de educación de requisitos.

Agradecimientos

Agradecemos al director de la Revista, Álvaro Rocha, por otorgar la oportunidad para dar a conocer los trabajos más relevantes presentados en el seminario SICC 2017. Asimismo, expresamos nuestra gratitud a los autores de los trabajos publicados, lo cual hizo posible este número en RISTI.

Referencias

- Biermann, A. W. (1997). *Great ideas in computer science: a gentle introduction*. Mit Press.
- Brookshear, G. G., & Brookshear, J. G. (2002). *Computer science: an overview*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.
- Denning, P. J. (2005). Is computer science science? *Communications of the ACM*, 48(4), 27–31.
- Henderson, H. (2009). *Encyclopedia of computer science and technology*. New York, NY, USA: Infobase Publishing.
- Newell, A., Perlis, A. J., & Simon, H. (1967). What is computer science. *Science*, 157(3711), 1373–1374.
- Ralston, A., Reilly, E. D., & Hemmendinger, D. (1976). *Encyclopedia of computer science*. New York, NY, USA: Petrocelli/Charter.

Índice / Index

EDITORIAL

Ciencias de la Computación en el mundo real.....	ix
<i>Gloria Piedad Gasca-Hurtado, Liliana Machuca-Villegas</i>	

ARTIGOS / ARTICULOS / ARTICLES

Desarrollo de un Analizador en línea, de las concentraciones de Cobre en soluciones acuosas del proceso de Extracción por Solvente, mediante procesamiento de imágenes.....	1
<i>Claudio Ayala B, José Gallardo A, Manuel Olivares A</i>	

Comparación de vocalizaciones de Espadarana prosoblepon en Dos Tipos de Ecosistemas usando Agrupamiento Jerárquico	13
<i>Diana C. Duque Montoya, Claudia V. Isaza Narváez, Estefany Cano Rojas</i>	

Evaluación de mapas de competencias educativas: una propuesta difusa basada en 2-tuplas.....	22
<i>Diego F. Duran, Gabriel E. Chanchí, Jose L. Arciniegas</i>	

Propuesta de una Arquitectura para Agricultura de Precisión Soportada en IoT	39
<i>Edwin Andrés Quiroga Montoya, Sergio Fernando Jaramillo Colorado, Wilmar Yesid Campo Muñoz, Gabriel Elías Chanchí Golondrino</i>	

Las redes sociales web como herramienta para gestionar información en procesos de co-creación.....	57
<i>Liliana González-Palacio, Mauricio González-Palacio, Jairo Ortiz-Pabón</i>	

Factores sociales y humanos que afectan el proceso de educación de requerimientos: una revisión sistemática	69
<i>Ítalo Donoso Barraza, Vianca Vega Zepeda</i>	

Desarrollo de un Analizador en línea, de las concentraciones de Cobre en soluciones acuosas del proceso de Extracción por Solvente, mediante procesamiento de imágenes

Claudio Ayala B.¹, José Gallardo A.², Manuel Olivares A.³

claudio.ayala@uantof.cl, jgallardo@ucn.cl, manuel.olivares@ucn.cl

¹ Departamento de Ingeniería Eléctrica, Universidad de Antofagasta, Av. Angamos 601, 1240000, Antofagasta, Chile.

^{2,3} Departamento de Ingeniería de Sistemas y Computación, Universidad Católica del Norte, Av. Angamos 0610, 1240000, Antofagasta, Chile.

DOI: 10.17013/risti.24.1-12

Resumen: Este artículo presenta el desarrollo de un nuevo analizador, instrumento para medir en línea concentraciones de cobre en acuoso. Su principio de funcionamiento se basa en la aplicación de técnicas de procesamiento de imágenes y un sistema automatizado para la toma de muestras desde el proceso. En principio, este analizador ha sido desarrollado para el proceso de extracción por solvente, que es una de las etapas del proceso hidrometalúrgico para la obtención de cátodos de cobre. Los resultados logrados son alentadores pues el instrumento desarrollado presenta una precisión aceptable para este tipo de instrumentos no convencionales, augurándose mediciones con mayor precisión cuando se implemente el instrumento en su versión industrial, con componentes de mayor calidad y robustez. Con este nuevo instrumento, las empresas mineras tendrán una alternativa de medición en línea de esta importante variable operacional, mejorando la operación de la planta y el incremento en su eficiencia y recuperación.

Palabras clave: Analizador, concentración, cobre, extracción, solvente.

Development of a copper concentration on-line Analyzer, for aqueous solutions from the Solvent Extraction process, through images processing and colorimetry

Abstract: This article proposes a new analyzer, instrument to measure on line copper concentrations in aqueous. Its operating principle is based on technics of images processing and colorimetry, and an automated system for the taking of samples directly withdrawn from the process. In principle, this analyzer has been developed for the solvent extraction process, which is one of the stages of the hydrometallurgical process to obtain copper cathodes. The results of this development are promising due to the analyzer presents an acceptable precision for this type of non conventional instruments, remaining even a space for improvements

when using components in its industrial version, of greater quality and robustness. With this new analyzer, the mining companies will have an alternative to measure on line this important operational variable, improving, with its application, the operation of the plant and increasing its efficiency and recovery.

Keywords: Analyzer, concentration, copper, extraction, solvent.

1. Introducción

El proceso hidrometalurgia del cobre, compuesto por las etapas de Lixiviación (LIX), Extracción por Solvente (SX) y Electro-obtención (EW) es uno de los más importantes métodos de producción de cobre desde minerales oxidados de baja ley (Komulainen, Doyle III, Rantala, & Jämsä-Jounela, 2009). La investigación en este proceso, ha sido enfocada principalmente en los procesos químicos y equipos asociados. Hasta el presente, el control de los procesos de SX-EW ha recaído sobre simples lazos de control y sintonía manual, con operadores manipulando activamente los set-point (SP) para controladores de nivel básico o de primer nivel (Komulainen et al., 2009; Bergha, Jämsä-Jounela, & Hodouinc, 2001). Un factor clave que reduce la eficiencia de plantas de SX es la metodología reactiva que los operadores de estas plantas están obligados a realizar, basados en análisis de laboratorio de los flujos fuera de línea, los cuales ya han sido procesados a través de la planta (Li, Yuan, Yang, & Huang, 2007). Los ingenieros de proceso utilizan estos análisis *After-the-fact* para predecir los mejores parámetros de operación para el subsecuente período de operación. Estos parámetros de operación predichos resultan en ajustes manuales para los parámetros del proceso, los cuales, si son correctos, facilitan una extracción más alta de cobre y rechazo más alto a las impurezas. El nivel de éxito de estas decisiones está sujeto a cómo el *staff* operacional interpreta el estado de la planta en un momento dado (Li et al., 2007).

Una de las razones de esta forma de operar plantas de SX es la falta de medidores de las variables operacionales más relevantes del proceso, como ser la Razón orgánico-acuoso (O/A), Tiempo de Separación de Fases (TSF) y Concentraciones de Cobre. Sin duda que la variable operacional más importante dentro del proceso de SX es la concentración de cobre en los flujos de salida de las distintas etapas del proceso, principalmente la concentración de cobre del electrolito rico, que sirve de entrada al proceso de electro-obtención. Existe por lo tanto la necesidad de medir las concentraciones de cobre en soluciones acuosas en tiempo real, con el objetivo de realizar el control de esta importante variable operacional, en plantas de extracción por solvente.

El principal objetivo de control de una planta de extracción por solvente de cobre es maximizar la recuperación y proveer electrolito rico estable para el proceso de electro-obtención (EW) (Komulainen et al., 2009). Una concentración adecuada de cobre en el electrolito rico permite una operación más estable y una operación más eficiente en la obtención de los cátodos de cobre, por estos motivos es muy importante mantener estable su valor operacional lo más cercano posible a las consignas establecidas por el proceso.

Sin embargo, la mayoría de las plantas de SX en Chile y en el mundo, no poseen o disponen de un medidor de concentraciones en línea, principalmente por el elevado costo de estos instrumentos. La mayoría de las plantas realizan medidas de concentraciones fuera de

línea, sobre muestras tomadas en distintos puntos del proceso, demorando dos o más horas en realizar un análisis de ella en un laboratorio.

Respecto a las alternativas existentes en el mercado de la instrumentación analítica, de medidores de concentración de cobre en línea para procesos mineros, solo existen los analizadores Courier. Este tipo de analizador, permite un muestreo y análisis en forma automática y consistente. La precisión de las mediciones es comparable a la de un laboratorio que analiza muestras manuales de alta calidad. Los analizadores Courier en línea utilizan la misma tecnología de análisis de Fluorescencia de Rayos X por Dispersión de Longitud de Onda (Martínez, Gil Novoa, & Barón González, 2006), que se utiliza en los analizadores de laboratorio de alto rendimiento. Sus principales desventajas son el alto costo y la necesidad de re-calibración permanente.

En este reporte, se presenta el desarrollo de un nuevo medidor de concentraciones de cobre en línea, basado en el procesamiento de imágenes y la técnica de colorimetría. Algoritmos genéticos son utilizados para la obtención de los parámetros que permitan la mejor estimación del modelo de correlación planteado, es decir, la estimación con el menor error cuadrático, respecto de un universo de muestras con diferentes valores de concentraciones de cobre obtenidas en laboratorio. Con posterioridad, este modelo de correlación es aplicado sobre muestras capturadas en línea directamente del proceso. La principal ventaja de este analizador en línea respecto a lo existente, es su relativo bajo costo.

Este artículo es estructurado como sigue: en la sesión 2 se describe el proceso de mezcla y decantación de la técnica de extracción por solventes y en la sesión 3 un esquema general del instrumento, con una descripción de cada una de sus partes y el principio de medición utilizado. En la sesión 4 se muestran los resultados experimentales y en las sesiones 5 y 6 se presentan las conclusiones, trabajo futuro y referencias bibliográficas respectivamente.

2. Proceso

En la Figura 1, se representa un estanque mezclador-decantador, típico de los procesos de extracción y re-extracción utilizados en plantas hidrometalurgias para la obtención de cátodos de cobre.

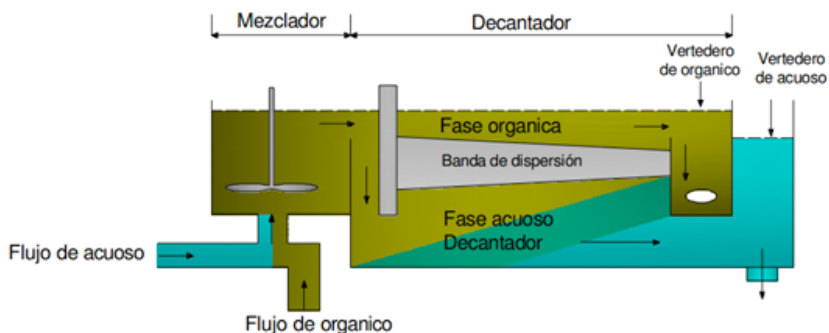


Figura 1 – Estanque mezclador- decantador para el proceso de extracción y re-extracción.

El orgánico y el acuoso que entran a los estanques principal y secundario son mezclados, produciéndose, durante esta acción, un traspaso de cobre del acuoso al orgánico (o de orgánico a acuoso en re-extracción). La medición se debe realizar en el vertedero de acuoso (salida de acuoso), para conocer la concentración de cobre producto del proceso de extracción (o re-extracción) que se ha producido en esta etapa.

3. Principio de medición y tecnología implementada

La tecnología propuesta se basa en la técnica de la colorimetría (Artigas, Capilla, Malo, Martínez, Luque, Pujol, 2002) aplicada a una muestra tomada desde el proceso, utilizando un algoritmo de procesamiento de imágenes. La toma de la muestra y su procesamiento es automatizado para la medición en línea, utilizando dispositivos como válvulas, sensores y un controlador lógico programable.

3.1. Esquema de medición

La Figura 2 representa un esquema del sistema para la medición en línea de la concentración de cobre.

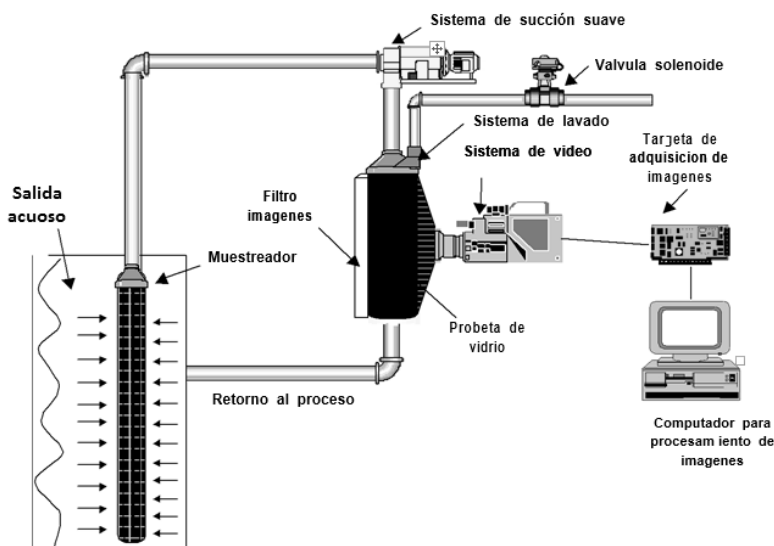


Figura 2 – Esquema del instrumento de medición

Mediante una bomba de succión, se extrae la muestra de acuoso en la salida de cada decantador, donde se desea obtener una medida de la concentración de cobre. Esta muestra se deposita en una probeta de vidrio, previamente sometida a un proceso de lavado, con agua controlada por la válvula solenoide. Luego en forma sincronizada, se toman imágenes de la muestra mediante una cámara y se procesa la información en un computador. Esta secuencia se vuelve a repetir periódicamente en un tiempo que puede ser fijado por el usuario.

3.2. Rocesamiento de las imágenes de las muestras

El proceso de estimación de la concentración de cobre en acuoso se realiza utilizando un simple algoritmo de tratamiento de las tomas (imágenes) captadas de las muestras que se extraen del acuoso. Estas imágenes son descompuestas en sus colores primarios Red, Green y Blues, modelo RGB (Rojo, Verde y Azul). Estos colores primarios, combinados de forma adecuada, conforman una gama amplia de colores, como por ejemplo las combinaciones que se ilustran en la Figura 3.

RGB es un modelo de color basado en la síntesis aditiva, con el cual es posible representar un color mediante la mezcla por adición de los tres colores de luz primarios (Li et al., 2007; Sankaranarayanan & Sivanantha, 2006). El modelo de color RGB, no define por sí mismo lo que significa exactamente rojo, verde o azul, por lo que los mismos valores RGB pueden mostrar colores notablemente diferentes en varios dispositivos que usen este modelo de color, y aunque utilicen un mismo modelo de color, sus espacios de color pueden variar considerablemente.



Figura 3 – Modelo de color RGB.

El modelo RGB asigna un valor numérico de 1 a 3 dígitos, que representa a cada uno de los colores primarios, el que a su vez determina la proporción en que es mezclado cada color, de manera que un valor “0” para cada valor de RGB significaría que éste no interviene en la mezcla y a medida que el valor aumenta, éste aporta más intensidad a la coloración. Así de manera usual, la intensidad de cada uno de los colores componentes, se mide según en una escala que va del 0 a 255 (Li et al., 2007; Sankaranarayanan & Sivanantha, 2006).

La técnica de procesamiento de imagen utilizada, se basa en que la solución contenedora del ion específico tiene un color definido en función de la concentración de cobre que posea, de manera que ésta solución puede ser correlacionada con una determinada concentración del ion (en nuestro caso ion Cobre II) específico. La metodología consiste entonces en utilizar el color de la solución como un medio para medir, a través de un programa de tratamiento de imágenes, la intensidad de los colores primarios y basados en esta medición, encontrar un modelo matemático de estimación que permita correlacionar la imagen de color obtenida, con la concentración de cobre en la solución.

A partir de las diversas imágenes capturadas para cada tipo de muestra desde el proceso, se aplica el procesamiento de imágenes que consiste en la descomposición de la imagen en sus canales Rojo, Verde y Azul (RGB). Como resultado, para cada imagen capturada a una frecuencia de muestreo definida, se obtienen tres matrices cuyas celdas son valores numéricos que varían entre 0 y 255. Con posterioridad, para cada una de las matrices obtenidas se determina la media y este valor numérico escalar, es asignado a las variables R, G y B. Los valores escalares de R, G y B, posteriormente son utilizados en un modelo matemático que permite estimar, en tiempo real, la concentración de cobre para cada una de las muestras capturadas. Los parámetros de la ecuación que modela la correlación de los componentes R, G y B de cada muestra con su concentración de cobre, fueron obtenidos mediante la aplicación de Algoritmos Genéticos, cuya función objetivo fue la minimización del error cuadrático entre los valores de concentración de cobre determinados mediante ensayos de laboratorio y el valor estimado por el modelo de correlación. En la siguiente sección se presenta este procedimiento, de manera más detallada.

4. Resultados experimentales

Para comprobar la reproducibilidad de los datos obtenidos al realizar mediciones, se toman múltiples muestras con diferentes concentraciones en un intervalo comprendido entre 30 g/L y 46,0 g/L de electrolito rico de Cobre (II). Estas muestras son preparadas por dilución a partir de la muestra de mayor concentración, como muestra la Figura 4. El valor de concentración de estas muestras, fue determinada en un Laboratorio Químico de la Universidad de Antofagasta, mediante el método de la espectrofotometría.

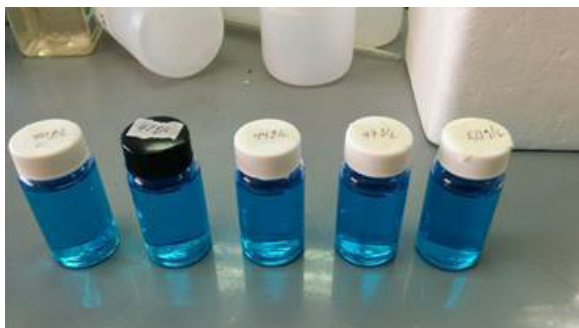


Figura 4 – Muestras obtenidas por dilución.

Las mediciones se realizan empezando con la muestra de menor concentración (se capturan diversas imágenes de cada muestra), para concluir con la muestra de mayor concentración, repitiendo la medición en varias ocasiones para observar qué comportamiento tiene la técnica en relación con los valores obtenidos y analizar la reproducibilidad de estas mediciones. Con los datos conseguidos, se genera la Tabla 1 (en esta tabla solamente se presenta solo un pequeño número de los experimentos realizados).

Como se puede observar en la Tabla 1, los datos (escalares) de los tres grupos, R, G y B, mantienen valores muy cercanos para cada muestra de una determinada concentración preparada en laboratorio.

<i>Experimento</i>	<i>Muestra</i>	<i>Concentración [g/l]</i>	<i>R</i>	<i>G</i>	<i>B</i>
1	1	30,00	20	121	158
2	1	30,00	21	123	160
3	1	30,00	21	121	158
4	1	30,00	20	123	160
1	2	35,00	18	118	158
2	2	35,00	19	119	158
3	2	35,00	18	118	158
4	2	35,00	18	115	158
1	3	41,30	17	117	161
2	3	41,30	17	115	158
3	3	41,30	17	116	160
4	3	41,30	17	116	160
1	4	42,20	16	116	161
2	4	42,20	16	117	162
3	4	42,20	16	116	160
4	4	42,20	16	115	160
1	5	43,10	15	115	162
2	5	43,10	16	116	163
3	5	43,10	15	113	159
4	5	43,10	14	113	161
1	6	44,50	14	112	161
2	6	44,50	14	116	165
3	6	44,50	14	116	166
4	6	44,50	13	114	165
1	7	46,00	13	110	166
2	7	46,00	12	108	163
3	7	46,00	12	108	163
4	7	46,00	12	107	163

Tabla 1 – Resultado de cuatro mediciones por cada muestra.

Como se mencionó en la sección anterior, haciendo un proceso de ajuste de datos basado en los resultados experimentales, se sigue un procedimiento para obtener un modelo que represente la concentración de cobre en función de los valores RGB.

Para trabajar con los modelos de estimación se utilizan: la ecuación de segundo grado 1.1 y la ecuación de primer grado 1.2.

$$Conc = A_1 * Fcol^2 + A_2 * Fcol + A_3 \quad (1.1)$$

$$Fcol = C_1 * R + C_2 * G + C_3 * B \quad (1.2)$$

Se tiene de esta manera un modelo que tiene 6 coeficientes que deben ser determinados. La determinación de estos coeficientes se realiza mediante un algoritmo de optimización, el cual minimiza el error cuadrático entre el valor de concentración medido y la concentración estimada mediante las ecuaciones (1.1) y (1.2). El modelo de optimización se presenta en (1.3).

$$\min x \sum_k^{35} (CON_K^{med} - CON_k^{estimada}(x))^2 \quad (1.3)$$

con $x^{min} \leq x \leq x^{max}$; CON_k^{med} corresponde a la concentración medida (en laboratorio) en el experimento k, $CON_k^{estimada}$ es la concentración estimada para el experimento k mediante las ecuaciones (1.1) y (1.2). El vector de parámetros de diseño queda definido por $x = [C_1 \ C_2 \ C_3 \ A_1 \ A_2 \ A_3]$.

Para resolver el problema de modelamiento se utilizaron y compararon diferentes técnicas clásicas de identificación, habiéndose elegido finalmente la técnica de Algoritmos Genéticos (GA) que entregó un menor error de ajuste.

Una vez determinados los parámetros de diseño, la ecuación (1.1) entregará la concentración estimada y representa el modelo matemático que permite asociar la tonalidad que exhibe la muestra con su concentración de cobre, a través de los coeficientes A_1 y $Fcol$ (Factor color). La ecuación (1.2) determina el $Fcol$ en función de los valores RGB obtenidos en el procesamiento de la imagen.

Para trabajar con los modelos propuestos, se usan los datos proporcionados por la Tabla 1, datos que fueron procesados mediante el Toolbox de Algoritmos Genéticos de Matlab, para generar los coeficientes C_1, C_2, C_3, A_1, A_2 y A_3 , que son mostrados en Tabla 2.

C1	C2	C3
0.94893858	0.21013730	0.33678871
A1	A2	A3
-0.09042402	-3.89993823	3.99293590

Tabla 2 – Coeficientes C_1, C_2, C_3, A_1, A_2 y A_3

Para el proceso de validación del modelo, se utilizan los coeficientes de la Tabla 2 y las ecuaciones (1.1) y (1.2). Los resultados se consignan en la Tabla 3, en la cual se registra el $Fcol$, la concentración estimada y el cálculo del error al cuadrado (SE) entre la concentración de la muestra y la concentración arrojada por el modelo. Con estos valores de SE, se determina el Error Cuadrático Medio, MSE, y la Raíz del Error Cuadrático Medio o RMSE, valores que se registran en la Tabla 4.

<i>Muestra</i>	<i>Conc. muestra [g/l]</i>	<i>Fcol</i>	<i>Conc. Modelo [g/l]</i>	<i>SE</i>
1	30,00	-8.807231	31.33	1.76
1	30,00	-8.133654	29.73	0.07
1	30,00	-8.111595	29.68	0.10
1	30,00	-8.723745	31.13	1.28
2	35,00	-11.335520	36.58	2.50
2	35,00	-11.335520	36.58	2.50
2	35,00	-11.672309	37.19	4.82
2	35,00	-11.965932	37.71	7.36
3	41,30	-13.841751	40.65	0.42
3	41,30	-13.588448	40.29	1.02
3	41,30	-13.378311	39.98	1.73
3	41,30	-13.715100	40.47	0.69
4	42,20	-15.127478	42.30	0.01
4	42,2	-13.653672	40.38	3.30
4	42,2	-14.453901	41.47	0.53
4	42,20	-15.210964	42.39	0.04
5	43,10	-15.486325	42.70	0.16
5	43,10	-15.337616	42.54	0.32
5	43,10	-15.906600	43.15	0.00
5	43,10	-16.518750	43.74	0.41
6	44,50	-18.749619	45.33	0.68
6	44,50	-17.572281	44.60	0.01
6	44,50	-16.898704	44.07	0.18
6	44,50	-17.235493	44.35	0.02
7	46,00	-19.782044	45.76	0.06
7	46,00	-21.973544	46.03	0.00
7	46,00	-21.151257	46.03	0.00
7	46,00	-20.814468	45.99	0.00

Tabla 3 – Resultado de concentraciones y SE determinadas a partir del modelo.

<i>Indicador</i>	<i>Valor</i>
MSE	1.18
RMSE	1.09

Tabla 4 – Indicadores estadísticos de los datos experimentales.

Finalmente, se configura el modelo en la unidad de procesamiento de imágenes y se realizan las lecturas para las concentraciones de cada una de las muestras.

La Tabla 5 registra las concentraciones de las muestras, el valor de concentración entregado por el analizador, correspondiente a cada muestra, y el valor del Fcol correspondiente.

Concentración muestra [g/l]	Fcol	Concentración medida [g/l]
30.00	-8.192188	29.73
35.00	-11.364503	36.58
41.30	-13.538357	40.29
42.20	-14.554653	41.47
43.10	-15.949196	43.15
44.50	-17.812161	44.60
46.00	-20.949184	45.99

Tabla 5 – Concentración de la muestra, Fcol y Concentración medida en g/l.

La Figura 5 muestra un gráfico de los valores registrados en la Tabla 5. Se observa un buen ajuste del modelo a los valores de las muestras, comportamiento esperado de acuerdo a los indicadores determinados previamente, MSE y RMSE.

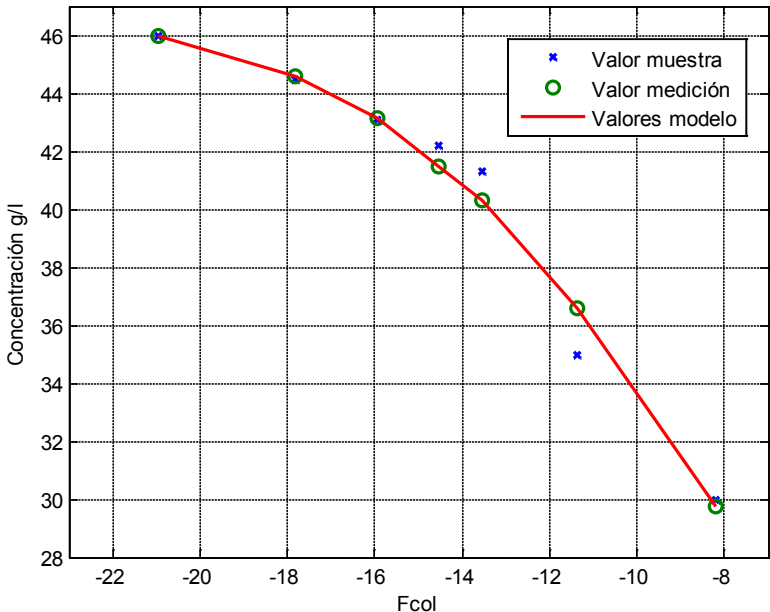


Figura 5 – Factor Color versus Concentración muestra, Concentración medida y Concentración modelo.

4.1 Equipos y programas utilizados para configurar el Analizador

- Computador con procesador Intel Core i3, 8 Gb RAM DDR,
- Tarjeta de adquisición de vídeo 3DEMON de 8MB,
- Cámara de vídeo, modelo SONY CCD-TRV34,
- Plataforma .NET: procesamiento de las imágenes,
- PLC RSLogic 1100: automatización del proceso de medición.

5. Conclusiones y trabajo futuro

De acuerdo a los resultados experimentales, se demuestra que existe una relación importante entre las concentraciones de soluciones acuosas y la combinación de los colores primarios del modelo RGB, en el rango de concentraciones estudiados, lo que motiva a seguir perfeccionando el analizador desarrollado. Se debe considerar además que por ser este instrumento un prototipo, fue implementado con componentes y recursos constructivos de muy bajo coste, por ejemplo, se utilizó una cámara del tipo webcam para la captura de imágenes, se puede esperar que la precisión de las mediciones mejore, utilizando componentes de una mejor calidad.

La relevancia de poder contar con un analizador de concentraciones en línea, y que sea de bajo costo, es alta para la operación y control del proceso de SX, siendo ésta su principal variable operacional.

Un aspecto importante a considerar para el perfeccionamiento del analizador, es que los cambios en tipo y cantidad de los elementos que componen las soluciones acuosas, en las diferentes plantas de SX, puede requerir compensaciones en el sensor para mantener su precisión. Una forma de compensar es recalibrando el equipo para cada planta de SX a cada cierto período, parámetro que se determinará en las pruebas a realizar en plantas industriales.

Como trabajo futuro, se propone sustituir el modelo de estimación representado por la ecuación 1.1, por una red neuronal de aprendizaje supervisado, que permita a partir de los componentes RGB de las imágenes capturadas, estimar la concentración de cobre que posee el acuoso, sustituyendo además el computador por una tarjeta embebida para el procesamiento de la red neuronal ya entrenada.

Finalmente, se debe destacar que esta innovación se encuentra en proceso de evaluación de la protección industrial en INAPI (Chile), con número de solicitud 2665-2016 (prioridad), con fecha 20/10/2016.

Referencias

- Artigas, J. M., Capilla P., Malo J., Martínez F., Luque M., & Pujol J., (2002), *Fundamentos de colorimetría*, Valencia, España: Universitat de València.
- Bergha, L.G., Jämsä-Jounela, S.L., & Hodouinc, D. (2001). State of the art in copper hydrometallurgic processes control Bergha. *Control Engineering Practice* 9, 1007–1012.
- Komulainen, T., Doyle III, F., Rantala, A., & Jämsä-Jounela, S. L. (2009). Control of an industrial copper solvent extraction process. *Journal of Process Control*, 19, 2–15.

- Li, G., Yuan, R., Yang, Z. & Huang, X. (2007). A yellow license plate location method based on RGB model of color image and texture of plate. *In Digital Media and its Application in Museum & Heritages, Second Workshop on* (pp. 42–46). IEEE.
- Martínez, B., Gil Novoa, O. D., & Barón González, A. J. (2006). Espectrometría de Fluorescencia de Rayos X. *Revista Colombiana de Física*, 38(2), 790–793.
- Mwale, M. & Megaw, D. (2011). Development of effective solvent-extraction process control – low cost implementation value-addition to hydrometallurgical copper operations. *In Proceedings Sixth Southern African Base Metals Conference. Southern African Institute of Mining & Metallurgy*, 353–365, Johannesburg.
- Sankaranarayanan, K., Sivanantha Raja, A., (2006). Use of RGB Color Sensor in Colorimeter for better Clinical measurement of Blood Glucose. *BIME Journal*, 6(1), 23–28.

Comparación de vocalizaciones de *Espadarana prosoblepon* en Dos Tipos de Ecosistemas usando Agrupamiento Jerárquico

Diana C. Duque Montoya ¹, Claudia V. Isaza Narváez ¹, Estefany Cano Rojas ²

diana.duquem@udea.edu.co , victoria.isaza@udea.edu.co , canoestefany@gmail.com

¹ SISTEMIC - Universidad de Antioquia, calle 70 No. 52 – 21, AA 1226, Medellín, Colombia.

² GHA - Universidad de Antioquia, calle 70 No. 52 – 21, AA 1226, Medellín, Colombia.

DOI: 10.17013/risti.24.13–21

Resumen: Se analizaron las vocalizaciones de la *Espadarana prosoblepon* en dos ecosistemas diferentes, un bosque conservado cerca de una represa en el suroeste de Antioquia - Colombia y un bosque alterado en Carmen de Atrato, Chocó - Colombia. Las grabaciones fueron segmentadas, luego procesadas para extraer y seleccionar los descriptores y finalmente clasificadas usando un algoritmo jerárquico. Con el fin de generalizar el análisis, también se procesaron las vocalizaciones de la *Sachatamia albomaculata*. Los resultados muestran que las vocalizaciones de la misma especie son afectadas por el ambiente, y que los tres conjuntos de vocalizaciones pueden ser diferenciados con la metodología propuesta.

Palabras clave: *Espadarana prosoblepon*, bioacústica, dendrograma, ecosistemas.

Comparison of Espadarana Prosoblepon Vocalizations in Two Different Ecosystems using Hierarchical Clustering

Abstract: *Espadarana prosoblepon* vocalizations were analyzed in two different ecosystems, a conserved forest near to a dam in southeast Antioquia and a disturbed forest in Carmen de Atrato, Chocó. The recordings were segmented, then processed to extract and select features and classified with a hierarchical algorithm. To analyze more generally the results, *Sachatamia albomaculata* vocalizations from the dam were processed too. The results show that the species vocalizations are clearly affected by the environment, and the three types of vocalizations can be separated with the proposed methodology.

Keywords: *Espadarana prosoblepon*, bioacoustics, dendrogram, ecosystems.

1. Introducción

La bioacústica es una ciencia que estudia el comportamiento de las especies mediante los sonidos que estas producen (Fletcher, 2007). Los beneficios de esta ciencia son la

posibilidad de estudiar especies muy raras que no son fáciles de encontrar en salidas de campo; el incremento del tiempo de estudio de las señales biológicas, por la capacidad de las grabadoras de funcionar por largos periodos de tiempo; y la poca afectación al ecosistema que causa al compararse con la recolección de ejemplares de las especies.

Actualmente, el GHA (Grupo Herpetológico de Antioquia) de la Universidad de Antioquia, trabaja en el monitoreo de anuros basados en sus vocalizaciones en el área que rodea la hidroeléctrica Jaguas, de Isagen (Bedoya, Isaza, Daza, & López, 2014; Daza, 2013). Los anuros son considerados una especie indicadora, porque son muy sensibles a los cambios en el ecosistema, por lo que pueden indicar la degradación del hábitat (Bedoya, 2015). Además, sus vocalizaciones tienen la ventaja de ser, en gran mayoría, monosilábicas y los cantos de la misma especie se encuentran por lo general, en la misma banda de frecuencia.

El objetivo de este estudio es analizar las diferencias en el espectrograma de los cantos de dos poblaciones de la misma especie de anuros, pero localizadas en diferentes lugares: Jaguas, una represa en el sureste de Antioquia, en Colombia y Carmen de Atrato, un pequeño pueblo del Chocó, también en Colombia. El primer hábitat es un bosque conservado, cuya única alteración es la represa construida cerca de este. El otro hábitat es un bosque intervenido por el ser humano. Para comparar los resultados y evaluar los cambios de los cantos de la especie, también se consideraron los cantos de otra especie de la misma familia. La metodología se basa en la descomposición en valores singulares de los descriptores extraídos, y en técnicas de agrupamiento jerárquico para obtener diferente cantidad de grupos por nivel de similitud.

2. Materiales y Métodos

2.1. Materiales

La especie escogida para este estudio es la *Espadarana prosoblepon*, un anuro de la familia *Centrolenidae*. Este anuro habita centro y suramérica, desde Honduras hasta Ecuador; es una especie nocturna y sus cantos tienen una frecuencia dominante alrededor de los 5 KHz (Richman, 2008). Las vocalizaciones de la *Sachatamia albomaculata* fueron usadas para comparar los resultados obtenidos, esta especie pertenece a la misma familia *Centrolenidae* pero la banda de frecuencia de sus cantos está alrededor de los 6KHz (Chantasirivisal, 2005). Las vocalizaciones de esta segunda especie se obtuvieron de las grabaciones de Jaguas únicamente.

Para este estudio se procesaron 58 grabaciones con cantos de las dos especies, 28 grabaciones de Jaguas, que fueron registradas en Noviembre del 2012 y 30 grabaciones de Carmen de Atrato, registradas en Octubre del 2015. Las señales se obtuvieron con micrófonos omnidireccionales usando una grabadora *Song Meter SM2*. Esta grabadora fue programada para grabar durante un minuto, con frecuencia de muestreo de 44100 Hz y guardar las grabaciones en formato WAC, que posteriormente fueron transformadas sin pérdida a formato WAV.

2.2. Métodos

En la propuesta realizada para el análisis de las grabaciones, primero se hace una segmentación de las vocalizaciones, es decir, la identificación y extracción de las partes de

las grabaciones con cantos de las ranas; estos segmentos fueron caracterizados en tiempo y frecuencia. El segundo paso es la extracción de características; dado que el sonido es solo una señal, se calcularon los coeficientes *cepstrales* de frecuencia para describir cada segmento usando diferentes variables; luego se transformaron estas variables para reducir su número mediante una técnica de filtrado basada en la descomposición en valores singulares; y por último se clasificaron los datos.

Segmentación

La segmentación se hizo de forma manual. Para llevarla a cabo, se inspeccionaron visualmente los espectrogramas de las señales y cuando un canto era reconocido, la banda de frecuencia y la ventana de tiempo se registraban para ser ingresados al extractor de características. En la figura 1, se presenta un ejemplo de los segmentos espectrales de ambas especies en los dos hábitats estudiados.

Extracción de Características

Para la extracción automática de características, se usó una versión modificada de los coeficientes *cepstrales* de Mel (MFCC). Los MFCC son muy utilizados para el reconocimiento automático del habla (Mermelstein, 1976). Sin embargo, esta técnica se usa para la identificación de vocalizaciones humanas, ya que se aplican los filtros de frecuencia de Mel que extraen coeficientes según el sistema auditivo humano, por lo que no es recomendable para las vocalizaciones de otras especies. Por lo tanto, el uso de estos filtros fue omitido durante el proceso, y de esta manera obtener coeficientes uniformemente y linealmente distribuidos a lo largo del espectro (Bedoya, 2015). El procedimiento fue el siguiente:

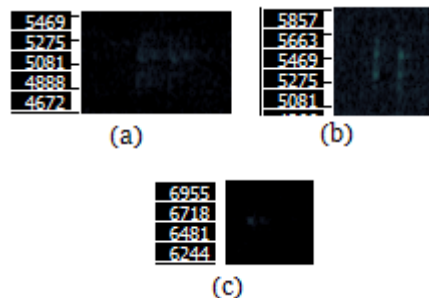


Figura 1 – Segmentos espectrales de los cantos de las ranas. A la derecha de cada segmento se muestra la escala de frecuencia en Hz. (a) Canto de la *Espadarana prosoblepon* en Jaguas. (b) Canto de la *Espadarana prosoblepon* en Carmen de Atrato. (c) Canto de la *Sachatamia albomaculata* en Jaguas.

1. Se divide cada segmento en partes.
2. Se calcula la potencia espectral de cada parte usando la transformada rápida de Fourier (FFT)

3. Se obtiene el logaritmo de la energía de todas las partes del segmento.
4. Por último, se estima la transformada de coseno discreta de los resultados para obtener los coeficientes *cepstrales*.

El número de coeficientes es determinado por el tipo de aplicación. Para este caso se usaron 100 coeficientes, basándose en un trabajo previo (Bedoya, 2015).

Selección de Características

De la etapa anterior se obtuvieron 100 variables por canto, lo que dificulta su procesamiento y visualización; por lo tanto, se optó por un método estadístico para reducir la dimensionalidad del problema. El escalamiento multidimensional (EMD) es un método basado en la descomposición en valores singulares (Golub & Kahan, 1965) para transformar las variables originales en un nuevo conjunto de variables que conserve la distancia entre las muestras (Kruskal, 1964). Se aplicó este método porque la técnica de agrupamiento a utilizar se basa en distancia, por lo que se esperan resultados aproximados a los que se habrían obtenido con el uso de las variables originales.

Este método comienza con el cálculo de la matriz de distancia euclidiana estandarizada $\mathbf{D}_{n \times n}$, donde n es el número de muestras y δ_{ij} representa a cada uno de sus elementos. Luego se obtienen las matrices $\mathbf{A}$ y $\mathbf{B}$ usando (1) y (2) respectivamente; a continuación, se extrae la matriz $\mathbf{Z}$ usando descomposición en valores singulares para obtener las nuevas variables como columnas de la nueva matriz (ver (3)). Los valores propios de la matriz $\mathbf{B}$ representan la significancia de las nuevas variables, es decir, las variables con los valores propios más altos son las más significativas para representar la muestra.

$$\mathbf{A}_{n \times n} = (a_{ij}) = \left(-\frac{1}{2} \delta_{ij}^2 \right) \quad (1)$$

$$\mathbf{B}_{n \times n} = (b_{ij}) = \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \mathbf{A} \left(\mathbf{I} - \frac{1}{n} \mathbf{J} \right) \quad (2)$$

$$\mathbf{B} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^T = \mathbf{V} \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{\Lambda}^{\frac{1}{2}} \mathbf{V}^T = \mathbf{Z} \mathbf{Z}^T \quad (3)$$

En (2) $\mathbf{I}$ es la matriz de identidad $n \times n$, $\mathbf{J}$ es una matriz $n \times n$ en donde todos sus elementos son uno; en (3) $\mathbf{V}$ y $\mathbf{\Lambda}$ son las matrices resultantes de la descomposición en valores singulares, y los elementos de la diagonal de $\mathbf{\Lambda}$ son las raíces cuadradas de los valores propios de $\mathbf{B}$ o el *stress* de las nuevas variables.

Agrupamiento

Se eligió el agrupamiento jerárquico para clasificar los datos (Jiawei & Kamber, 2001). De esta manera se dejan en un mismo grupo los cantos que se consideran comunes. Las variables más significativas obtenidas del paso anterior se usaron para calcular la distancia entre muestras, luego estas son agrupadas por su proximidad; se inicia con

grupos de un solo elemento hasta agrupar todos los elementos en un solo grupo. Lo que es interesante es analizar las clases que se forman durante este proceso para estudiar los diferentes niveles de similitud entre las muestras. La medida de similitud fue la distancia euclidiana estandarizada, como en el paso anterior, y las muestras fueron conectadas usando enlace simple (entre vecinos más cercanos) y enlace promedio para comparar.

3. Resultados

De la etapa de segmentación, 91 cantos de la *Espadarana prosoblepon* y 54 cantos de la *Sachatamia albomaculata* fueron encontrados en las grabaciones de Jaguas; y 63 cantos de la *Espadarana prosoblepon* fueron encontrados en las grabaciones de Carmen de Atrato. Los cantos de la *Espadarana prosoblepon* estaban todos en la misma banda de frecuencia, aunque los cantos de Carmen de Atrato fueron más difíciles de detectar debido al ruido presente en las grabaciones, al parecer porque la grabadora se encontraba cerca de un arroyo y la lluvia era frecuente. Fue interesante notar que las vocalizaciones de las ranas de Jaguas se encontraron en la franja horaria entre la 1 y las 5 de la mañana, mientras que las de Carmen de Atrato se encontraron entre las 6 y 10 de la noche.

Luego de la etapa de extracción de características, el escalamiento multidimensional fue ejecutado. El *stress* de todos los valores propios encontrados se muestra en la figura 2.

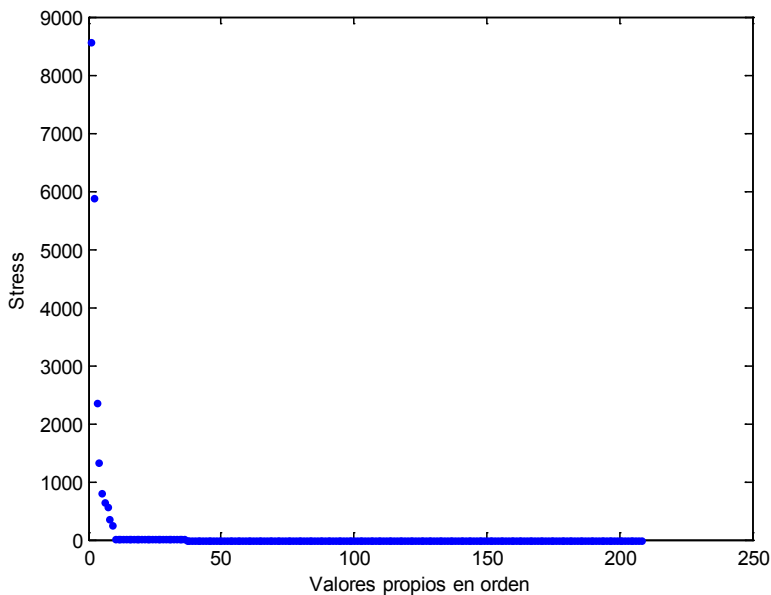


Figura 2 – *Stress* obtenido del escalamiento multidimensional.

La figura 2 muestra que las primeras nueve variables son las más significativas, lo que quiere decir que al calcular la distancia entre estas nueve nuevas variables es casi lo

mismo que calcular la distancia con las 100 variables originales. Estos fueron entonces los descriptores escogidos para realizar el agrupamiento.

En la figura 3 se grafica la variable 1 versus la variable 2 para todas las muestras. Los puntos azules son las vocalizaciones de la *Espadarana prosoblepon* de Carmen de Atrato, los puntos rojos son las vocalizaciones de la misma especie en Jaguas, y los puntos verdes son las vocalizaciones de la *Sachatamia albomaculata*. Es fácil identificar que las dos poblaciones de la misma especie se pueden separar, lo que quiere decir que las vocalizaciones de Jaguas son diferentes de las vocalizaciones del Chocó para la misma especie. Sin embargo, la diferencia con las vocalizaciones de la otra especie no es tan grande como se esperaba, puede ser incluso posible que la clase de la *Espadarana prosoblepon* del Carmen de Atrato se agrupe primero con la clase de la *Sachatamia albomaculata* que con la clase de la *Espadarana prosoblepon* de Jaguas; esto se describe más adelante. Como comentario adicional, las varianzas de las vocalizaciones de Jaguas parecen ser mayores a la de Carmen de Atrato, debido probablemente a que las ranas de Jaguas tienen una banda de frecuencias más amplia para cantar, mientras que las ranas en Carmen de Atrato estaban más limitadas por el ruido, por lo que el nicho acústico era más reducido.

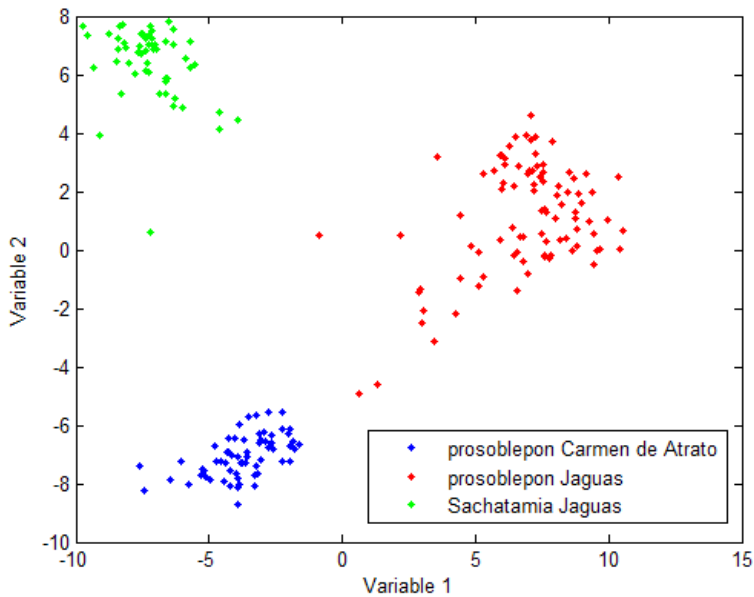


Figura 3 – Las muestras son graficadas en un espacio de dos dimensiones formado por la variable 1 y la variable 2 obtenidas del EMD. Los puntos azules son las vocalizaciones de la *prosoblepon* de Carmen de Atrato, los puntos rojos son las vocalizaciones de la *prosoblepon* de Jaguas y los puntos verdes son vocalizaciones de la *Sachatamia*.

Finalmente, se utiliza el agrupamiento jerárquico con el propósito de encontrar las similitudes dentro y entre especies. El dendrograma de enlace simple se muestra en la figura 4.

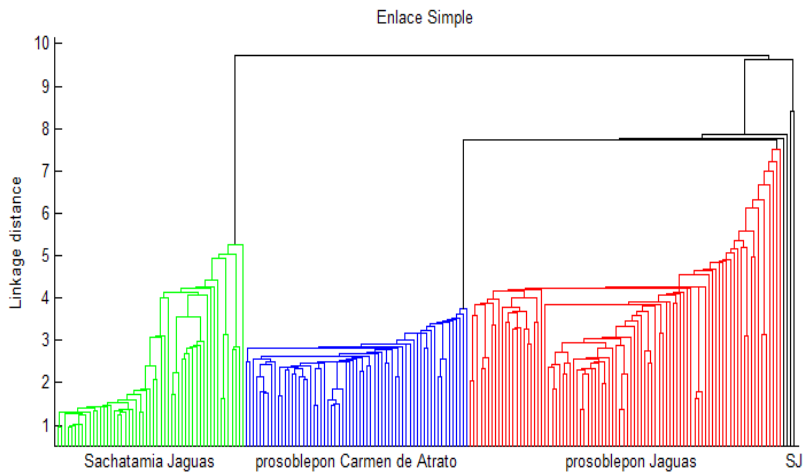


Figura 4 – Dendrograma de enlace simple de los datos. Los colores son los grupos encontrados automáticamente correspondientes a las tres poblaciones de ranas, las ramas negras son datos atípicos.

Las muestras de la *Espadarana prosoblepon* en Carmen de Atrato coincidieron con las ramas azules, las muestras de la misma especie de Jaguas coincidieron con las ramas rojas (excepto por tres muestras atípicas), y las muestras de la *Sachatamia albomaculata* coincidieron con las ramas verdes, excepto por un dato atípico. Las diferencias en las varianzas de los grupos explican la razón por la que las vocalizaciones de Carmen de Atrato son agrupadas primero en el dendrograma. Para un monitoreo biológico puede ser interesante analizar los subgrupos de cada clase para extraer información de esos individuos, especialmente para las muestras de Jaguas. También se pueden ver los datos atípicos como ramas negras, los cuales representan 4 errores del algoritmo al compararse con las etiquetas originales, el error fue entonces del 1.92%

También se implementó un dendrograma de enlace promedio para disminuir el efecto de los datos atípicos. Este se muestra en la figura 5.

En este caso, las muestras de *Sachatamia* y Carmen de Atrato se agrupan en una clase antes de agruparse con las muestras de la *prosoblepon* de Jaguas. Esto demuestra que la influencia de los factores ambientales puede causar mayor similitud entre vocalizaciones de diferente especie que entre vocalizaciones de la misma especie, pero en hábitat diferente. Comparando con las etiquetas originales, obtenidas por un análisis individual realizado por herpetólogos, el clasificador propuesto solamente tuvo dos errores, una muestra de *Sachatamia* y una de la *prosoblepon* de Jaguas clasificadas ambas como *prosoblepon* de Carmen de Atrato lo que corresponde a un error del 0.96%.

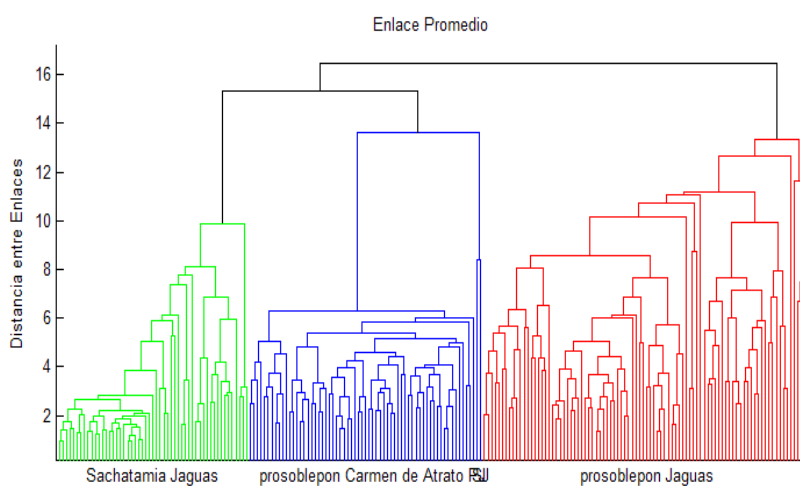


Figura 5 – Dendrograma de enlace promedio de los datos. Los colores son los grupos encontrados automáticamente correspondientes a las tres poblaciones de ranas.

4. Conclusiones

La metodología utilizada fue apropiada para el problema, ya que los grupos pudieron ser fácilmente separados usando distancia euclidiana estandarizada como medida de similitud; además, el escalamiento multidimensional ayudó a reducir el número de variables con poca pérdida de información de los datos originales.

Este estudio prueba que el monitoreo acústico puede detectar la influencia que el ambiente causa en las vocalizaciones de una misma especie. Para un estudio relacionado con el monitoreo bioacústico, con la propuesta realizada se pueden percibir cambios en el comportamiento de los anuros y analizar los subgrupos más representativos para estudiar las características fonéticas de las especies. Es interesante notar que en ninguno de los dos dendrogramas las muestras de Jaguas se agruparon primero que las de Carmen de Atrato; este hecho refuerza la hipótesis del nicho acústico, que supone que las especies diferentes dentro de un mismo lugar intentan diferenciar sus vocalizaciones para mejorar la comunicación entre los individuos de la misma especie (Krause, 1993).

5. Agradecimientos

Gracias al Grupo Herpetológico de Antioquia por el suministro de la base de datos y a Isagen por permitir el monitoreo de los ecosistemas alrededor de la hidroeléctrica Jaguas.

Referencias

Bedoya, C., Isaza, C., Daza, J. M., & López, J. D. (2014). Automatic recognition of anuran species based on syllable identification. *Ecological Informatics*, 24, 200–209. <http://doi.org/10.1016/j.ecoinf.2014.08.009>.

- Bedoya, C. L. (2015). *Unsupervised Identification of Anuran Vocalizations*. (Master's thesis). Medellín, Colômbia: Universidad de Antioquia.
- Chantasirivisal, P. (2005). *AmphibiaWeb - Sachatamia albomaculata*. Retrieved from http://amphibiaweb.org/cgi/amphib_query?where-genus=Sachatamia&where-species=albomaculata.
- Daza, J. M. (2013). *Monitoreo de algunas poblaciones de vertebrados endémicos y amenazados presentes en las Centrales Hidroeléctricas Jaguas y San Carlos en el Oriente Antioqueño* (Rep. No. 46/4208). Medellín: Universidad de Antioquia.
- Fletcher, N. (2007). Animal Bioacoustics. In T. D. Rossing (Ed.), *Springer Handbook of Acoustics* (1st ed.), pp. 473–491. New York: Springer New York. http://doi.org/10.1007/978-0-387-30425-0_19.
- Golub, G., & Kahan, W. (1965). Calculating the Singular Values and Pseudo-Inverse of a Matrix. *Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis*, 2, 205–224. <http://doi.org/10.1137/0702016>.
- Han, J., Kamber, M., & Pei, J. (2012). *Data mining: concepts and techniques* (3rd ed). Waltham, MA: Elsevier.
- Krause, B. L. (1993). The Niche Hypothesis: A Virtual Symphony of Animal Sounds, The Origins of Musical Expression and the Health of Habitats. *The Soundscape Newsletter*, 6, 6–10.
- Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1), 1–27. <http://doi.org/10.1007/BF02289565>.
- Mermelstein, P. (1976). Distance measures for speech recognition, psychological and instrumental. *Pattern Recognition and Artificial Intelligence*, 116, 374–388.
- Richman, S. (2008). *AmphibiaWeb - Espadarana Prosoblepon*. Retrieved from <http://amphibiaweb.org/species/1741>.

Evaluación de mapas de competencias educativas: una propuesta difusa basada en 2-tuplas

Diego F. Duran ¹, Gabriel E. Chanchí ², Jose L. Arciniegas ¹

dduran@unicauca.edu.co, gchanchi@unimayor.edu.co, jlarci@unicauca.edu.co

¹ Universidad del Cauca, Cll. 5 #4-70, 190003, Popayán, Colombia.

² Institución universitaria Colegio Mayor del Cauca, Cra. 7 #2-34, 190003, Popayán, Colombia.

DOI: 10.17013/risti.24.22-38

Resumen: Aunque el uso de términos lingüísticos permite a los docentes realizar una evaluación de competencias educativas de una forma natural y significativa, la evaluación de mapas de competencias puede ser una tarea imprecisa, ya que debe ser considerada una gran cantidad de información sobre las relaciones existentes entre elementos (competencias y actividades de evaluación), pesos de tales relaciones y niveles de competencia. Con el propósito de abordar este problema, este trabajo propone un modelo difuso basado en 2-tuplas lingüísticas para evaluar competencias educativas teniendo en cuenta las características de los mapas de competencias. Adicionalmente, se presenta una aplicación Java, la cual fue usada para evaluar la propuesta mediante datos reales. Los resultados obtenidos permiten concluir que la propuesta es efectiva y eficiente.

Palabras-clave: Educación, Evaluación, Lógica difusa, Mapas de competencias.

Competency maps assessment: a fuzzy proposal based on 2-tuples

Abstract: Although the use of linguistic terms enables teachers carry out educational competencies assessment in a natural and significant way, the competency maps assessment may be an imprecise task because teachers have to consider a large quantity of information related to relations among elements (i.e., competencies and assessment activities), weights of such relations and proficiency levels. In order to address this issue, this work proposes a fuzzy model for assessing competencies based on the features of competency maps by using the 2-tuples approach. Additionally, a Java application is presented, which was used to assess the proposal by means of real data. The results indicate that the proposal is quite effective and efficient.

Keywords: Assessment, Competency maps, Education, Fuzzy Logic.

1. Introducción

En el contexto educativo, las competencias son las diferentes habilidades que los estudiantes pueden aprender y los docentes enseñar. Éstas son de gran importancia

porque definen criterios públicos de la calidad educativa de una institución (OCDE, 2006). Ejemplos de tal importancia son la educación basada en competencias, la cual es desarrollada en algunos países de América Latina (por ejemplo, Colombia y Perú) y la Declaración de Bolonia, en la cual el concepto de evaluación de competencias fue seleccionado como el parámetro básico para comparar la educación superior entre diferentes universidades europeas (en 29 países como España, Alemania y Francia) (Serrano, Romero & Olivas, 2013). En esta dinámica, la evaluación de competencias es una tarea clave. Los docentes califican Actividades de Evaluación (AE) asignando Niveles de Competencia (NC) de acuerdo al desempeño del estudiante. Esta tarea puede ser realizada usando notas numéricas en el rango [0,10]; sin embargo, el uso de términos lingüísticos (por ejemplo, Bajo, Medio, Alto) puede ser más natural y significativo en la evaluación de algunos tipos de AE y competencias (Fung, 2013), por ejemplo, si estas involucran habilidades y capacidades observables de los estudiantes en actividades de solución de problemas.

Aunque el uso de términos lingüísticos es útil, la evaluación de competencias puede ser un proceso difícil e impreciso cuando el docente evalúa estructuras en las que se establecen relaciones entre ellas, ya que debe ser considerada una gran cantidad de información. Éste es el caso de los mapas de competencias, los cuales sirven como herramientas para observar los NC alcanzados en el currículo de un curso o programa (García & Gairín, 2011). De acuerdo con el borrador de estándar *Simple Reusable Competency Map* (SRCM), los mapas de competencias están compuestos por competencias, relaciones de tipo general-específica entre éstas con etiquetas de pesos (estos ponderan la importancia de una competencia específica en el desarrollo de una general), y los NC asignados en los procesos de calificación. Con el propósito de abordar el problema mencionado, el presente artículo propone un enfoque difuso para evaluar competencias, basado en la estructura y características de los mapas de competencias usando el enfoque de 2-tuplas lingüísticas. El enfoque difuso es usado porque permite una mejor representación del lenguaje natural (Herrera, López, Alonso, Moreno, Cabrerizo & Porcel, 2009). Adicionalmente, en este trabajo se desarrolló una herramienta Java basada en la propuesta, la cual se encarga de computar los NC de todas las competencias de los mapas mientras mide y presenta los tiempos de ejecución utilizados.

El resto del artículo está organizado como sigue. En la Sección 2 se presentan los conceptos introductorios y los trabajos previos relacionados con esta investigación. La Sección 3 describe el enfoque propuesto, incluyendo la construcción de mapas de competencias y del perfil del estudiante (desde el punto de vista computacional) y el proceso de cómputo de los NC. La Sección 4 presenta una evaluación de la propuesta en términos de efectividad y tiempos de ejecución, utilizando la herramienta Java desarrollada y datos reales provistos por los docentes. Finalmente, en la Sección 5 se presentan las conclusiones y trabajos futuros derivados de la presente investigación.

2. Aspectos preliminares

Esta sección presenta un grupo de conceptos, aspectos y trabajos relacionados con la presente investigación. Dentro de estos se encuentran: los mapas de competencias y los problemas existentes en su evaluación, los enfoques difusos, incluyendo los basados en 2-tuplas y en el principio de extensión.

2.1. Mapas de competencias

Los mapas de competencias son herramientas usadas por los docentes para observar los NC alcanzados por los estudiantes en el currículo de una materia, área, programa o curso (García & Gairín, 2011). Una buena aproximación a la estructura de un mapa de competencias es dada por el borrador de estándar SRCM (Ostyn, 2006). Éste define un modelo de información simple y reusable para mapas de competencias, para ser usado en la descripción, referencia e intercambio de datos que definen relaciones entre competencias en el contexto educativo. En el modelo de SRCM, un Grafo Acíclico Dirigido (GAD) es usado para describir los mapas, en el cual cada competencia es representada por un nodo. Dado que un GAD es una colección jerárquica de nodos, una competencia puede ser descompuesta en sub-competencias, o que tales sub-competencias contribuyen a ella. La Figura 1 muestra un fragmento de un mapa de competencias basado en el modelo de SRCM, el cual fue provisto por un docente de matemáticas que usa los estándares Básicos de Competencias de Colombia para establecer las metas de enseñanza y aprendizaje (Ministerio de Educación de Colombia, 2006). Como puede observarse, los mapas tienen las siguientes características:

- Las competencias tienen sub-elementos que pueden ser AE, e incluso, otras competencias (notar que una competencia puede ser un sub-elemento o sub-competencia de otra). Entonces, los mapas pueden estar compuestos por uno o más niveles.
- Las AE diseñadas por los docentes (exámenes orales o escritos, solución de problemas, etc.) han sido incluidas como sub-elementos de algunas competencias. En la práctica educativa, éstas son usadas por los docentes para realizar un proceso de calificación de los estudiantes, en el cual unos NC son asignados.
- La estructura establece relaciones entre sus elementos, las cuales son etiquetadas con un peso que representa la importancia de una sub-competencia en el desarrollo de una competencia de un nivel superior.

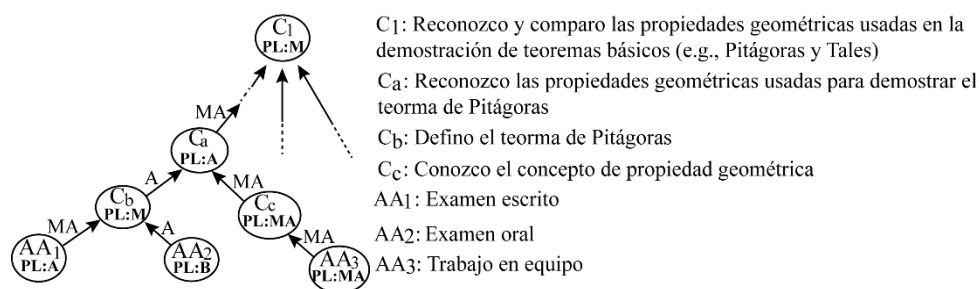


Figura 1 – Fragmento de un mapa de competencias basado en el modelo de SRCM

- Los NC han sido asignados por el docente a todos los elementos del mapa. Los docentes usualmente usan los NC asignados a los AE para computar mentalmente los de todas las competencias.

En el ejemplo presentado, pesos y NC fueron asignados de acuerdo a un grupo de siete términos lingüísticos como el siguiente:

$$S = \{N : Nulo, MB : \text{Muy Bajo}, B : \text{Bajo}, M : \text{Medio}, A : \text{Alto}, MA : \text{Muy Alto}, T : \text{Total}\} \quad (1)$$

Como comentario final, aunque la misma estructura puede ser usada para varios estudiantes, cada mapa de competencia corresponde a un solo estudiante porque registra sus propios NC.

2.2. Problemas en la evaluación de competencias

De acuerdo con (Fung, 2013), el criterio del docente en un proceso de evaluación en un AE está usualmente determinado por grados numéricos de 1 a 5, o términos lingüísticos asociados con números, como en el caso de las escalas Likert (Likert, 1932), por ejemplo, “Muy Bajo” está asociado a 1, “Bajo” a 2, “Medio” a 3, “Alto” a 4, “Muy Alto” a 5. Cuando el docente observa estas posibles opciones, su cerebro procesa información del proceso de evaluación y elige una para asignar un NC que describa su valor de percepción. Sin embargo, existen algunos aspectos que pueden entorpecer este proceso, tales como la incertidumbre, restricciones, y el vago conocimiento que pueda tener el docente. Incluso, es muy difícil para él asegurar que los criterios numéricos sean igualmente aplicados al evaluar a todos los estudiantes. De acuerdo con (Montero, Alias, Badía, Fonseca & Vicent, 2014), estos problemas se dan especialmente si la AE involucra habilidades y capacidades observables en actividades de solución de problemas (por ejemplo, una mesa redonda, experimentos químicos, etc.) o cuando las competencias evaluadas son de una naturaleza subjetiva, tales como la competencia “trabajo en equipo”. Con el propósito de resolver estos problemas, en trabajos como (Fung, 2013), (Zhaoa & Yueb, 2012) y (Massanet, Riera, Torrens & Herrera, 2014), la evaluación es realizada usando términos lingüísticos como los mostrados en la expresión (1), los cuales hacen el juicio del docente (y de cualquier experto en determinado contexto) más confiable y consistente.

Asumiendo que los docentes usan términos lingüísticos para llevar a cabo procesos de evaluación de AE, la evaluación de mapas puede ser una tarea imprecisa debido a que éstos deben considerar una gran cantidad de información lingüística. En consecuencia, los NC pueden ser determinados incorrectamente, especialmente, en los niveles superiores de los mapas.

2.3. Enfoque lingüístico difuso

Hay situaciones en las cuales no es apropiado el uso de una evaluación cuantitativa. En estos casos, la utilización de un enfoque lingüístico puede ser más conveniente, en el cual, son usados un conjunto de términos como los mostrados en la expresión (1). De acuerdo con (Herrera & Martínez, 2000), una forma de generar un conjunto lingüísticos consiste en distribuir los términos en una escala con un orden bien establecido. Un ejemplo es $S = \{s_0 : N, s_1 : MB, s_2 : B, s_3 : M, s_4 : A, s_5 : MA, s_6 : T\}$. Como puede observarse, un grupo de índices han sido incluidos en la expresión (1). De acuerdo con el criterio de expertos que participaron en esta investigación, éste es el conjunto más apropiado en la práctica educativa en cuanto a términos y granularidad (esto es, la cardinalidad del conjunto de términos).

Por otro lado, en este trabajo son usadas las funciones de pertenencia triangulares (ver la Figura 2). La semántica de los siete términos se define de la siguiente forma:

$N=(0,0,0.17)$, $MB=(0,0.17,0.33)$, $B=(0.17,0.33,0.5)$, $M=(0.33,0.5,0.67)$,
 $A=(0.5,0.67,0.83)$, $MA=(0.67,0.83,1)$, $T=(0.83,1,1)$

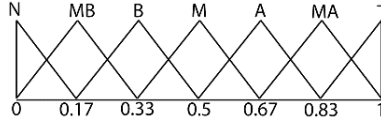


Figura 2 – Un grupo de siete términos lingüísticos y su semántica

Dentro del enfoque lingüístico difuso, se encuentran enfoques computacionales como el basado en 2-tuplas y el basado en el principio de extensión, los cuales son importantes para la presente propuesta.

2.4. Enfoque computacional basado en 2-tuplas

Este enfoque está basado en el concepto de translación simbólica, la cual se define como sigue: sea un conjunto de términos lingüísticos, si un método simbólico agregando información lingüística obtiene un valor $\beta \in [0, g]$, y $\beta \notin \{0, \dots, g\}$ entonces una función de agregación ($app_2(.)$) es usada para expresar el índice del resultado en S . La translación simbólica de un término lingüístico s_i es un valor numérico evaluado en $[-0.5, 0.5]$ que apoya la “diferencia de información” entre una cantidad de información $\beta \in [0, g]$ obtenida después de una operación de agregación simbólica y el valor más cercano en $\{0, \dots, g\}$ que indica el índice del término lingüístico más cercano en $S (i = round(\beta))$. Este enfoque representa la información lingüística por medio de 2-tuplas (s_i, α_i) , $s_i \in S$ y $\alpha_i \in [-0.5, 0.5]$, siendo s_i una representación de la etiqueta lingüística centro de la información y α_i un valor numérico que expresa el valor de la translación desde el resultado original β a la etiqueta del índice más cercana i en el conjunto de términos lingüísticos s_i , esto es, la translación simbólica. Además, se define un conjunto de funciones de transformación entre términos lingüísticos y 2-tuplas, y entre valores numéricos y 2-tuplas. Las 2-tuplas que expresan la información equivalente a β es obtenida con las siguientes funciones:

$$\Delta: [0, g] \rightarrow S \times [-0.5, 0.5] \quad (2)$$

$$\Delta(\beta) = \begin{cases} s_i, & i = round(\beta) \\ \alpha = \beta - i, & \alpha \in [-0.5, 0.5] \end{cases} \quad (3)$$

Donde $round(.)$ es la operación de redondeo común, s_i tiene la etiqueta más cercana a β y α es el valor de la translación simbólica. Como un comentario importante, la conversión de un término lingüístico a una 2-tupla consiste en añadir un valor cero de la siguiente forma:

$$s_i \in S \Rightarrow (s_i, 0) \quad (4)$$

2.5. Enfoque computacional basado en el principio de extensión

El principio de extensión ha sido introducido para generalizar operaciones matemáticas numéricas a los conjuntos difusos. Para esto, se usan operadores de agregación lingüísticos, los cuales actúan de acuerdo con:

$$S^n \xrightarrow{\tilde{F}} F(\mathcal{R}) \xrightarrow{app_1(\cdot)} S \quad (4)$$

Donde S^n simboliza el producto cartesiano n de S , $\tilde{F}$ es un operador de agregación basado en el principio de extensión, $F(\mathcal{R})$ es el conjunto de conjuntos difusos sobre el conjunto de números reales R , $app_1: F(\mathcal{R}) \rightarrow S$ es una función de aproximación lingüística que retorna una etiqueta del conjunto de términos lingüísticos cuyo significado es el más cercano al número difuso no etiquetado obtenido y S es el conjunto de términos inicial.

2.6. Trabajos relacionados

Existen trabajos en la literatura centrados en evaluar competencias usando enfoques lingüísticos difusos. En (Pengshun & Jianguo, 2012) se presenta una metodología para evaluar competencias laborales de los estudiantes. Tal propuesta aplica un método de evaluación difuso basado en una estructura de competencias profesionales, la cual es una estructura en árbol de tres niveles con competencias pre-definidas. En (Serrano, Romero & Olivas, 2013) se propone un enfoque difuso para recomendar actividades educativas. Inicialmente, se define un modelo basado en 2-tuplas que posibilita la evaluación lingüística de competencias sobre una estructura compuesta de un grupo de competencias y actividades. La estructura define que cada competencia está relacionada con un grupo de actividades. Después, se define un modelo de recomendación que usa un filtro basado en los resultados de la evaluación. En (Zhaojun, Xiao & Xiaoqing, 2010) se presenta un modelo para evaluar competencias educativas basado en un enfoque difuso. En un caso de estudio, el modelo es aplicado a un sistema de índices de competencias para vendedores, el cual es una estructura en árbol compuesta por 3 niveles. En (Jia & Zhao, 2010) se propone un método de evaluación de competencias básicas para empresas. Éste se basa en la entropía y los números difusos triangulares, posibilitando una evaluación objetiva de empresas. Éste puede ser útil en un proceso de decisión, en el cual un grupo de empresas son comparadas entre sí de acuerdo al criterio de expertos. En (Zhang, 2012) se propone un método de decisión basado en operadores de agregación. Éste usa el modelo de 2-tuplas para agregar los desempeños de las alternativas medidas por un grupo de evaluadores. El resultado de la agregación es usado para tomar una decisión del grupo de alternativas.

Todos los trabajos relacionados revelan la importancia de la evaluación de competencias y desempeños por medio de términos lingüísticos. Para este propósito, varios modelos y enfoques han sido propuestos para facilitar el proceso de evaluación con un alto nivel de efectividad. Sin embargo, no hay evidencia de investigaciones basadas en la estructura y características de los mapas de competencias (ver la Sección 2.2). Adicionalmente, aunque todas las propuestas son efectivas, no ha sido realizada una evaluación de su viabilidad en términos de eficiencia.

3. Propuesta de evaluación de competencias

La presente propuesta computa los NC de todos los elementos de los mapas de competencias de forma cualitativa, a partir de los alcanzados por los estudiantes en las AE. Para este propósito, el modelo considera las características de los mapas, con respecto a su estructura, tipos de elementos y pesos (ver la Sección 2.1). Las próximas sub-secciones inicialmente describen metodológicamente cómo son construidos los mapas de competencias y los perfiles de los usuarios desde el punto de vista computacional. Posteriormente, se describe el proceso basado en 2-tuplas para el cómputo de los NC de todo el mapa.

3.1. Fase A: Representación de la estructura del mapa de competencias

Los currículos educativos están basados en un grupo de competencias generales establecidas por un ente gubernamental (por ejemplo, el Ministerio de Educación Nacional a través de los Estándares básicos de competencias) o por la dirección general de una institución (por ejemplo, una escuela o universidad). Tales competencias deben ser alcanzadas por los estudiantes durante el ciclo escolar. Debido a la universalidad inherente a las competencias generales, a partir de éstas pueden definirse un grupo de sub-competencias, relacionadas unas con otras con diferentes pesos, de tal forma que al final se consigue establecer una guía para los docentes de cómo alcanzar las metas educativas. A partir de esta dinámica se definen unos vectores que en conjunto constituyen la estructura de los mapas:

1. Dado que una competencia está relacionada con sus sub-elementos, los cuales pueden ser otras competencias o AE según sea el caso, entonces una competencia se define por medio del vector $\{C_1, C_2, \dots, C_j, \dots, C_k\}$.
2. Dado que los pesos están asociados a la relación entre C_i relacionados y su sub-elemento C_j , éstos son representados por el vector $\{P_{i1}, P_{i2}, \dots, P_{ij}, \dots, P_{ik}\}$.

En este punto, la estructura del mapa puede ser representada por medio de los vectores descritos. Sin embargo, aún falta incorporar los NC del estudiante antes de iniciar el proceso de cómputo. Esta situación es tratada en las siguientes sub-secciones.

3.2. Fase B: Construcción del perfil del estudiante

Cuando un estudiante realiza una AE, el docente califica su desempeño asignando un NC lingüístico. Por consiguiente, el perfil del estudiante está compuesto por un grupo de NC alcanzados en cada AE AE_j o competencia C_j . Entonces, el grupo de NC correspondientes a los sub-elementos de C_j es representado por el vector $\{NC_1, NC_2, \dots, NC_j, \dots, NC_k\}$. En ciertos momentos del ciclo escolar es necesario conocer las competencias alcanzadas o adquiridas por el estudiante. Para este propósito, en este trabajo se propone un proceso de cómputo basado en 2-tuplas, el cual es expuesto en las próximas sub-secciones.

Ejemplo 1: construcción de vectores. La Figura 3 presenta un ejemplo de un mapa de competencias de tres niveles. De acuerdo al modelo de vectores descrito anteriormente,

la estructura del mapa se representa por medio de los vectores $C_1=\{C_2, C_3\}$, $C_2=\{AE_1, AE_2\}$, $C_3=\{AE_2, AE_3\}$, $P_1=\{A, MA\}$, $P_2=\{M, MA\}$ y $P_3=\{A, B\}$. Adicionalmente, el perfil del estudiante es representado por los vectores $NC_2=\{B, A\}$ y $NC_3=\{A, MA\}$.

3.3. Fase C: Cómputo de los Niveles de Competencia

Con el propósito de computar los NC de todos los elementos del mapa, el modelo propuesto agrega los NC del perfil del estudiante considerando los vectores de la estructura del mapa. Todas las sub-fases del proceso, las cuales son una adaptación del modelo basado en 2-tuplas, son explicadas a continuación.

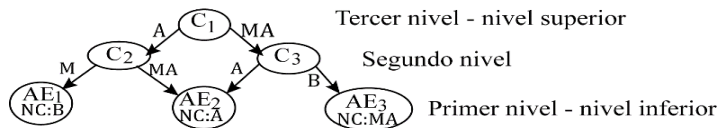


Figura 3 – Ejemplo de un mapa de competencias

Sub-fase C-1: Construcción de matrices

En esta sub-fase los vectores de las fases A y B son utilizados para construir el modelo matricial mostrado en la Figura 4, el cual posee las características pertinentes para la aplicación de operadores difusos de agregación. Este par de matrices son construidas por cada nivel del mapa. La primera, llamada Matriz de Niveles de Competencia (MNC), organiza las competencias en las columnas y sus sub-elementos en las filas. Los elementos de MNC son los NC NC_{ij} alcanzados por el estudiante en la AE_i o C_j , siendo estos sub-elementos de C_i . La segunda matriz, llamada Matriz de Pesos (MP), también organiza las competencias en las columnas y las AE o competencias en las filas. Sin embargo, en este caso cada elemento de MP corresponde al peso P_{ij} de la relación entre AE_j o C_j y C_i . En ambos casos, si no existe una relación entre una columna y una fila, el valor del elemento correspondiente es cero. En la Figura 4 se puede apreciar que las matrices organizan las AE o las competencias según sea necesario.

Matriz de Niveles de Competencia					Matriz de Pesos				
	C_1	C_2	...	C_n		C_1	C_2	...	C_n
$AE_1 - C_1$	NC_{11}	NC_{21}	...	NC_{n1}	$AE_1 - C_1$	P_{11}	P_{21}	...	P_{n1}
$AE_2 - C_2$	NC_{12}	...	...	...	$AE_2 - C_2$	P_{12}	...	...	...
...	...	...	...	...	...	...	...	...	...
$AE_k - C_k$	NC_{1k}	NC_{2k}	...	NC_{nk}	$AE_k - C_k$	P_{1k}	P_{2k}	...	P_{nk}

Figura 4 – Matrices de Niveles de Competencia y Pesos

Por otra parte, de acuerdo con el procesamiento de palabras llevado a cabo en la lógica difusa (ver Secciones 2.4 y 2.5), se da que en la agregación se obtiene inicialmente un valor numérico, el cual es posteriormente analizado para asignarle un valor lingüístico

(el cual es considerado como resultado final de la agregación). Para diferenciarlos, en lo que sigue de este trabajo son nombrados como Valor de Agregación Numérico (VAN) y valor de agregación lingüístico (VAL). Considerando que los NC computados para cierto nivel influyen en el cómputo de los NC de niveles superiores, en el presente enfoque se proponen dos posibilidades para realizar la agregación:

- Posibilidad P1. La agregación en todos los niveles es realizada usando los VAL del nivel inmediatamente inferior. En este caso, MNC y MP se compone de términos lingüísticos.
- Posibilidad P2. La agregación a partir del tercer nivel es realizada usando los VAN del nivel inmediatamente inferior (notar que en la agregación del segundo nivel son usados los NC del primer nivel, los cuales son términos lingüísticos). En este caso, MNC se compone de valores numéricos mientras que PM lo hace de términos lingüísticos.

El uso de una u otra posibilidad depende del criterio del usuario. De cualquier forma, ambas tienen implicaciones en el modelo, las cuales son tratadas en las siguientes sub-fases.

Sub-fase C-2: Conversión de términos lingüísticos a 2-tuplas

En esta sub-fase, cada término lingüístico que hace parte de las matrices es convertido a 2-tuplas. De acuerdo con la expresión (4), un valor de cero debe ser añadido a cada término con el propósito de hacer translaciones simbólicas. En caso de hacer uso de la segunda de las posibilidades presentadas en la sub-fase C-1, la conversión es aplicada únicamente sobre la MP.

Ejemplo 2: construcción de matrices. Considerando el Ejemplo 1, la Figura 5 presenta las matrices MNC y MP para el segundo nivel del mapa. Notar que todos los términos lingüísticos han sido convertidos a 2-tuplas.

MNC		MP			
	C_2	C_3		C_2	C_3
AE ₁	(L,0)	0	AE ₁	(M,0)	0
AE ₂	(H,0)	(H,0)	AE ₂	(VH,0)	(H,0)
AE ₃	0	(VH,0)	AE ₃	0	(L,0)

Figura 5 – MNC y MP para el Ejemplo 2.

Sub-fase C-3: Agregación

Una vez organizados los NC, ya sea de las AE o de las competencias de cierto nivel del mapa, se computan los NC de todos los elementos del nivel inmediatamente superior. Para esto, se proponen dos diferentes procesos de agregación para cada una de las posibilidades establecidas en la sub-fase C-1.

Agregación para la posibilidad P1. Se establece el uso del Operador de Media Ponderada (WAO por su significado en inglés) para 2-tuplas descrito en (Herrera & Martínez, 2000):

$$\bar{x}^e = \Delta \left(\frac{\sum_{i=1}^n \beta_i * w_i}{\sum_{i=1}^n w_i} \right) \quad (5)$$

Este operador permite a diferentes valores de β tener diferentes importancias en la naturaleza de la variable x . Por lo tanto, cada valor de x_i tiene un peso asociado w_i que indica su importancia en la naturaleza de la variable. En términos de competencia, este operador permite agregar diferentes NC alcanzados en AE o competencias considerando su importancia para la competencia de la cual son sub-elementos. La función Δ es tratada en la Sección 2.4. Los valores de β para todos los índices de los términos lingüísticos son definidos en la Figura 6. Por ejemplo, los valores de β para B, A y MA son 2, 4 y 5 respectivamente.

N	MB	B	M	A	MA	T
0	1	2	3	4	5	6

Figura 6 – Distribución de índices para los términos lingüísticos

Agregación para la posibilidad P2. Se establece el uso del operador WAO definido en la expresión (5) con una variación de la naturaleza de β . En este caso, al aplicar el WAO para cierta competencia, los valores de β usados corresponden a los VAN computados para sus sub-elementos.

Ejemplo 3: agregación de 2-tuplas. Considerando el ejemplo 2, en este ejemplo son calculados los NC para las competencias C_2 y C_3 usando la expresión (5):

$$NC_{C_2} = \Delta \left(\frac{2 * 3 + 4 * 5}{3 + 5} \right) = \Delta(3.25)$$

Siguiendo un proceso similar, se obtiene $NC_{C_3} = \Delta(4.33)$. La función Δ es utilizada para asignar términos lingüísticos en la siguiente sub-fase.

Sub-fase C-4: Asignación del término lingüístico

En esta sub-fase, la función Δ interpreta los VAN para asignar los correspondientes términos lingüísticos. De acuerdo con el modelo basado en 2-tuplas, la función Δ genera un par de 2-tuplas (S_a, α_a) y (S_b, α_b) que corresponden al término más cercano inferior y superior del VAN. S_a es asignado si $|\alpha_a| < |\alpha_b|$, mientras que S_b es asignado si $|\alpha_b| < |\alpha_a|$.

Ejemplo 4: agregación y asignación de términos lingüísticos. Teniendo en cuenta que la competencia C_1 está en el tercer nivel del mapa, el cómputo de su NC puede ser realizado de acuerdo a las dos posibilidades establecidas anteriormente:

Considerando con la posibilidad P1. La Figura 7 presenta las matrices para el nivel superior del mapa. NCM está compuesta por los VAL computados para C_2 y C_3 .

Entonces el NC para C_1 es calculado como sigue:

$$NC_{C_1} = \Delta \left(\frac{3 * 4 + 4 * 5}{4 + 5} \right) = \Delta(3.55)$$

Considerando la posibilidad P2. La Figura 8 muestra las matrices para el nivel superior del mapa. NCM está compuesta por los VAN computados para C_2 y C_3 . Entonces el NC para C_1 es calculado como sigue:

$$NC_{C_1} = \Delta \left(\frac{3.25 * 4 + 4.33 * 5}{4 + 5} \right) = \Delta(3.85)$$

MNC	MP
C_1	C_1
$C_2 \begin{bmatrix} (M,0) \end{bmatrix}$	$C_2 \begin{bmatrix} (H,0) \end{bmatrix}$
$C_3 \begin{bmatrix} (H,0) \end{bmatrix}$	$C_3 \begin{bmatrix} (VH,0) \end{bmatrix}$

Figura 7 – MNC y MP para el nivel superior usando la posibilidad P1

MNC	MP
C_1	C_1
$C_2 \begin{bmatrix} 3.25 \end{bmatrix}$	$C_2 \begin{bmatrix} (H,0) \end{bmatrix}$
$C_3 \begin{bmatrix} 4.33 \end{bmatrix}$	$C_3 \begin{bmatrix} (VH,0) \end{bmatrix}$

Figura 8 – MNC y MP para el nivel superior usando la posibilidad P2

Para la asignación de los términos lingüísticos, se tiene que $NC_{C_2} = \Delta(3.25)$. Dado que 3.25 está ubicado entre los términos M y A, la función da como resultado las 2-tuplas $(M, 0.25)$ y $(M, -0.75)$. Como $|0.25| < |-0.75|$, entonces el término asignado a C_2 es M. En un proceso similar se da que el término asignado para C_3 y C_1 es A en ambos casos. Sin embargo, notar que los VAN para C_1 fueron 3.55 y 3.85 para cada una de las posibilidades. Aunque en ambos casos el término asignado es A, el primer valor se ubica en un punto medio entre M y A, mientras que el segundo es claramente cercano a A. Esta diferencia de resultados se da porque usando la primera posibilidad hay una pérdida de información nivel por nivel, resultado de la función de redondeo inherente a la función Δ . En cambio, con el uso de la segunda posibilidad, el cómputo de los NC de un nivel considera todos los NC de los niveles inferiores, por lo cual no hay pérdida de información. Así, es probable que en mapas complejos (con más elementos y niveles), el término lingüístico asignado difiera entre una y otra posibilidad. Aunque son claras las consecuencias dadas en el uso de cada una, la elección de una de las posibilidades dependerá del criterio del docente.

4. Evaluación

Esta sección presenta inicialmente un grupo de experimentos realizados para evaluar la propuesta usando datos reales. Posteriormente se presentan los resultados obtenidos.

4.1. Visión general

Los experimentos diseñados para dar respuesta a las siguientes preguntas de investigación: (1) ¿Cuál es la efectividad de la propuesta?; (2) ¿Qué ventajas provee la propuesta en términos de tiempo de ejecución en comparación con el enfoque basado en el principio de extensión?

Con el propósito de responderlas, se desarrolló una herramienta software soportada en el lenguaje Java, la cual fue útil para realizar los experimentos. Ésta ofrece dos modos de operación para computar los NC de las competencias (ver Figura 9): el Modo 1 está basado en la presente propuesta y el Modo 2 está basado en el principio de extensión. Los datos que describen la estructura de los mapas y los NC de los estudiantes han sido adaptados al formato JSON y almacenados en una base de datos NoSQL MongoDB. La aplicación computa los NC de todos los elementos de los mapas, presentando tanto los VAL como los VAN. Adicionalmente, presenta los tiempos de ejecución medidos para cada mapa. Como puede observarse, la aplicación provee los datos necesarios para evaluar la propuesta.

4.2. Métricas

La métrica definida en la expresión (6) fue usada para dar respuesta a la primera pregunta de investigación, la cual se define como el porcentaje de NC dados por la aplicación que son iguales a los proveídos por los docentes. También, puede ser considerada como el porcentaje de aciertos de la propuesta.

$$A = \frac{N_{hits}}{T} * 100\% \quad (6)$$

Donde N_{hits} es el total de aciertos y T es el total de NC evaluados. Con respecto a la segunda pregunta de investigación, el tiempo de ejecución de la aplicación es una variable que depende de varios factores, tales como los recursos propios del computador o las tareas paralelas realizadas por él al realizar los experimentos. Con el propósito de reducir la variabilidad de las medidas del tiempo de ejecución, se ha calculado un valor promedio t_a de cien medidas para cada experimento a , de acuerdo con:

$$t_a = \frac{\sum_{i=1}^{100} t_{ia}}{100} \quad (7)$$

Donde t_{ia} es la i -ésima medida de tiempo tomada para el experimento a .

4.3. Experimentos y resultados

Con el propósito de evaluar la presente propuesta, se hizo uso de los datos de 75 mapas, los cuales fueron contruidos y provistos por un grupo de docentes basándose en los Estándares Básicos de Competencias en Matemáticas para Colombia. A fin de explicar

de forma clara los experimentos y resultados alcanzados, se ha decidido en adelante diferenciar lo que se constituye como la *estructura del mapa* de lo que representa el *mapa de competencias* en sí. Con *estructura del mapa* se refiere al grupo de competencias, sus relaciones y pesos, mientras que el *mapa de competencias* adiciona a la *estructura del mapa* los NC alcanzados por cada estudiante en las AE y competencias. Así, un grupo de mapas de competencias (uno por estudiante) puede ser el resultado de evaluar una misma estructura. Los 75 mapas usados en la evaluación obedecen a 15 diferentes estructuras, esto es, 5 mapas por estructura. Estos incluyen los NC alcanzados por un grupo de estudiantes en todas las AE y competencias del mapa, razón por la cual son propicios para dar cuenta de la efectividad de la propuesta.

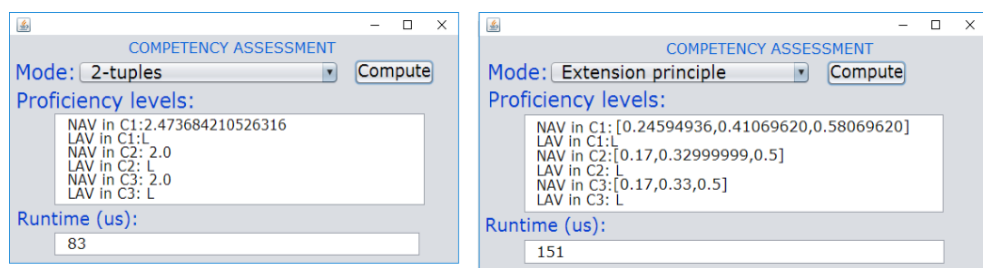


Figura 9 – Interfaces de la aplicación. A la izquierda se presenta la interfaz del modo 1 y a la derecha la del modo 2

Los experimentos y la colección de datos fueron realizados y obtenidos ejecutando las siguientes fases:

1. La estructura de los mapas fue adaptada al formato JSON y almacenada en la base de datos MongoDB. Cada estructura fue identificada por medio de un identificador único.
2. Los NC alcanzados en las AE fueron introducidos en la base de datos. Cada estudiante fue identificado con un identificador único.
3. Los NC para cada competencia (745 en total) de los mapas fue computada por la aplicación (usando el Modo 1 de la herramienta).
4. La fase 3 fue realizada utilizando el modelo basado en el principio de extensión (usando el Modo 2 de la herramienta).
5. Con el propósito de medir la efectividad de la propuesta, los NC dados por la aplicación en las fases 3 y 4 fueron comparados con los provistos por los docentes. Entonces, el porcentaje de aciertos fue medido mediante la expresión (6).
6. Con el propósito de conocer las ventajas de la presente propuesta en términos de tiempo de ejecución en comparación con el enfoque basado en el principio de extensión, todos los mapas fueron computados 100 veces para medir el tiempo de ejecución promedio de cada estructura y enfoque mediante la expresión (7). Cabe mencionar que los tiempos medidos fueron tomados para los métodos de la aplicación que están directamente relacionados a los procesos propios de cada enfoque (otros métodos como los relacionados con la gestión de bases de datos, interfaces, etc., no fueron considerados en las medidas).

La Tabla 1 presenta los resultados alcanzados en los experimentos de la fase 5. Los datos están organizados de acuerdo con cada uno de los dos enfoques evaluados y para cada estructura de los mapas. Analizando los resultados, el porcentaje de aciertos es 98.39% y 96.91% respectivamente. Por consiguiente, la efectividad del modelo puede ser considerada como positiva. Para validar la significancia estadística de los resultados, se quiere comprobar si existe una diferencia significativa entre los resultados de los dos enfoques. Inicialmente se plantean las siguientes hipótesis:

Estructura	Competencias por estructura	T	Propuesta			Enfoque basado en el principio de extensión	
			N_{hits}	$A_1(\%)$		N_{hits}	$A_2(\%)$
1	8	40	39	97.5		36	90
2	6	30	28	93.33		27	90
3	15	75	72	96		71	94.66
4	7	35	34	97.14		32	91.42
5	11	55	54	98.18		53	96.36
6	3	15	15	100		15	100
7	11	55	55	100		55	100
8	10	50	50	100		50	100
9	24	120	119	99.16		119	99.16
10	18	90	89	98.88		88	97.77
11	3	15	15	100		15	100
12	10	50	49	98		49	98
13	3	15	15	100		15	100
14	9	45	44	97.77		43	95.55
15	11	55	55	100		54	98.18
Total	149	745	733	98.39		722	96.91

Tabla 1 – Resultado de aciertos

H_0 (Hipótesis nula): no hay diferencias significativas entre los dos enfoques; H_a (Hipótesis alternativa): si existen diferencias significativas. Calculando una probabilidad de acierto promedio entre los enfoques se obtiene:

$$p = \frac{A_1 + A_2}{2} = \frac{0.9839 + 0.9691}{2} = 0.976655$$

Considerando un nivel de significancia $\alpha=0.05$, el valor estadístico crítico en la distribución normal es $Z_{\alpha=0.05}=1.96$. El error estándar es:

$$es = \sqrt{p(1-p)\left(\frac{1}{T} + \frac{1}{T'}\right)} = 0.007823$$

Para concluir acerca de las hipótesis, se evalúa si $|A_1 - A_2| = 0.184$ es mayor o menor que $es^*Z_{\alpha=0.05} = 0.0163$. Dado que $|A_1 - A_2|$ es mayor que $es^*Z_{\alpha=0.05}$, se concluye que la diferencia entre los aciertos del enfoque propuesto y los aciertos del basado en el principio de extensión es estadísticamente significativa. Por consiguiente, se rechaza H_0 y se acepta H_a como probablemente cierta. Por otro lado, las medidas de tiempo de ejecución promedio para cada enfoque y estructura son presentadas en la Figura 10. Analizando los resultados, los tiempos de ejecución para el enfoque basado en el principio de extensión son mayores que los de la propuesta entre una y dos veces en el 40% de los casos, y más de dos veces en el 60% de los casos. Éstos resultados son acordes con la complejidad computacional inherente a cada enfoque. Mientras que en la propuesta el proceso computacional se realiza sobre un único valor β que representa la cantidad de información de cada término lingüístico (gracias al uso de 2-tuplas), el enfoque basado en el principio de extensión obliga al uso constante de los 3 valores que constituyen la semántica de cada término. Esto presume un mayor consumo de memoria y procesador, lo cual explica las diferencias medidas en tiempos de ejecución entre cada enfoque.

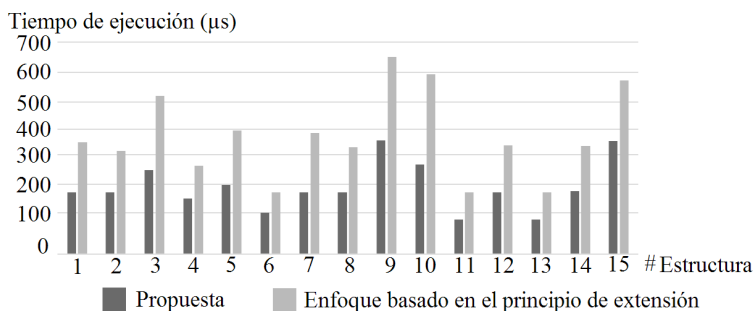


Figura 10 – Resultados de tiempos de ejecución

5. Conclusiones y trabajos futuros

La evolución de los sistemas computacionales implica cambios en el diseño y evaluación de los procesos de enseñanza y aprendizaje. En este caso, se ha propuesto un enfoque difuso para evaluar competencias, basado en las características de los mapas de competencias. El enfoque computa los NC de todos los elementos del mapa usando los alcanzados en las AE, permitiendo a estudiantes y docentes conocer las metas educativas alcanzadas en el currículo de un programa, área o curso. El enfoque ha sido implementado como una herramienta software Java y evaluado con datos reales de mapas provistos por docentes. Los resultados indican que es bastante efectivo, ofreciendo un porcentaje de aciertos cercano al 100% y más alto que el del enfoque basado en el principio de extensión. Además, provee valores bajos de tiempo de ejecución, por lo cual, es viable para ser integrado a plataformas de e-learning o servicios de internet para apoyar procesos educativos reales.

En trabajos futuros será estudiada la forma para incorporar la presente propuesta a un sistema de recomendación de contenidos de VoD (Video bajo Demanda). Con el propósito de proveer perfiles de usuario, los NC lingüísticos pueden ser adaptados

a una estructura de información (por ejemplo, matrices). Por lo tanto, el proceso de filtrado podría estar enfocado en las competencias no alcanzadas por los estudiantes. Adicionalmente, se buscará relacionar los contenidos con los mapas de competencias. En este sentido, el diseño de perfiles de aplicación de metadatos puede ser útil, adaptando diferentes esquemas como *TV-Anytime*, *Learning Object Metadata* (LOM), *Reusable Competency Definition* (RCD) y *Simple Reusable Competency Map* (SRCM).

Agradecimientos

Este trabajo es apoyado por el programa Doctorados Nacionales de Colciencias, convocatoria 617-2013.

Referencias

- Fung, K. K. (2013). A Fuzzy Qualitative Evaluation System: A multi-granular aggregation approach using fuzzy compound linguistic variable. *Journal of Intelligent & Fuzzy Systems*, 24, 61–78. doi:10.3233/IFS-2012-0531.
- García, M., & Gairín, J. (2011). Los Mapas de Competencias: una Herramienta para mejorar la Calidad de la Formación Universitaria. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 9(1), 84–102.
- Herrera, E., López, A., Alonso, S., Moreno, J., Cabrerizo, F., & Porcel, C. (2009). A computer-supported learning system to help teachers to teach fuzzy information retrieval systems. *Information Retrieval*, 12, 179–200. doi:10.1007/s10791-008-9087-3.
- Herrera, F., & Martínez, L. (2000). A 2-Tuple Fuzzy Linguistic Representation Model for Computing with Words. *IEEE Transactions on Fuzzy Systems*, 8(6), 746–752. doi:10.1109/91.890332.
- Jia, Z., & Zhao, L. (2010). Evaluation of Enterprise Core Competence Based on the Entropy and Triangular Fuzzy Number. In *Proceeding of the International Conference on E-Business and E-Government (ICEE)*, 2045–2048. doi:10.1109/ICEE.2010.517.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 1–55.
- Massanet, S., Riera, J. V., Torrens, J., & Herrera, E. (2014). A new linguistic computational model based on discrete fuzzy numbers for computing with words. *Information Sciences*, 248, 277–290. doi:10.1016/j.ins.2013.06.055.
- Ministerio de Educación de Colombia. (2006). Estándares básicos de competencias. Retrieved from http://www.mineducacion.gov.co/1621/articles-340021_recurso_1.pdf.
- Montero, J., Alias, F., Badía, D., Fonseca, D., & Vicent, L. (2014). A method for designing automatic assessment systems based on teachers reasoning for evaluating subjective engineering student's competences. In *Proceedings of the 9th Iberian Information Systems and Technologies (CISTI)*. doi:10.1109/CISTI.2014.6877046.

- OCDE. (2006). El Programa PISA de la OCDE. Retrieved from <http://www.oecd.org/pisa/39730818.pdf>.
- Ostyn, C., (2006). Simple Reusable Competency Map. Retrieved from <http://studylib.net/doc/18849165/simple-reusable-competency-map>.
- Pengshun, Z., & Jianguo, H. (2012). Application of Fuzzy Hierarchy Evaluation in the Occupational Competency Assessment of the Vocational Colleges. In *Proceedings of the 2012 International Symposium on Information Technology in Medicine and Education*, 182–185. doi:10.1109/ITiME.2012.6291276.
- Serrano, J., Romero, F., & Olivas, J. (2013). Hiperion: A fuzzy approach for recommending educational activities based on the acquisition of competences. *Information Sciences*, 248, 114–129. doi:10.1016/j.ins.2013.06.009.
- Zhang, H. (2012). The multiattribute group decision making method based on aggregation operators with interval-valued 2-tuple linguistic information. *Mathematical and Computer Modelling*, 56(1), 27–35. doi:10.1016/j.mcm.2012.01.001.
- Zhaoa, X., & Yueb, W. (2012). A multi-subsystem fuzzy DEA model with its application in mutual funds management companies' competence evaluation. *Procedia Computer Science*, 1, 2469–2478. doi:10.1016/j.procs.2010.04.279.
- Zhaojun, Z., Xiao, Q., & Xiaoqing, L. (2010). Competency Assessment Study Based on IPCA-Fuzzy Comprehensive Evaluation. In *Proceedings of the International Conference on Management and Service Science (MASS)*. doi:10.1109/ICMSS.2010.5578053.

Propuesta de una Arquitectura para Agricultura de Precisión Soportada en IoT

Edwin Andrés Quiroga Montoya¹, Sergio Fernando Jaramillo Colorado¹, Wilmar Yesid Campo Muñoz¹, Gabriel Elías Chanchí Golondrino²

eaquirogam@uqvirtual.edu.co, sfjaramilloc@uqvirtual.edu.co, wycampo@uniquindio.edu.co, gchanchi@unimayor.edu.co

¹ Universidad del Quindío, Cra. 15 Cll. 12 norte, 630004, Armenia, Colombia.

² Institución universitaria Colegio Mayor del Cauca, Cra. 7 #2-34, 190003, Popayán, Colombia.

DOI: 10.17013/risti.24.39-56

Resumen: Mediante este artículo se presenta una propuesta de arquitectura para la agricultura de precisión soportada en tecnologías abiertas de Internet de las Cosas (IoT) para la obtención, monitoreo y análisis de variables climáticas. La propuesta se basa en la estructura de la arquitectura Lambda, considerando diferentes capas como son: La capa de captura de datos, cuya función es la obtención de variables asociadas a un cultivo. La capa de almacenamiento, cuya función es recopilar la información en tiempo real desde los sensores. La capa de procesamiento, genera predicciones y recomendaciones, la cual es evaluada mediante pruebas de carga con el fin de determinar su capacidad y tiempo de respuesta. La capa de consulta, permite a los usuarios finales visualizar en una interfaz web los datos climáticos y las predicciones. Así, esta arquitectura pretende servir de referencia para la implementación de servicios basados en IoT en el área de la agricultura.

Palabras clave: Agricultura de Precisión, Arquitectura Lambda, Internet de las Cosas, Sparkjava Framework, Variables climáticas.

Proposal of an Architecture for Precision Agriculture Supported in IoT

Abstract: This paper presents a proposal for an architecture for precision agriculture supported in open technologies of the Internet of things (IoT) for the collection, monitoring and analysis of climatic variables. The proposal is based on the structure of the Lambda architecture, considering different layers such as: The data capture layer, whose function is to obtain variables associated with a crop. The storage layer, whose function is to collect the information in real time from the sensors. The processing layer generates predictions and recommendations, which are evaluated by load tests to determine their capacity and response time. The query layer allows to end users to view weather data and predictions on a web interface. Thus, this architecture aims to serve as a reference for the implementation of services based on IoT in the area of agriculture.

Keywords: Climate variables, Internet of Things, Architecture Lambda, Precision Agriculture, Sparkjava Framework.

1. Introducción

Actualmente, en el área de la agricultura una gran cantidad de datos son obtenidos sobre el suelo y las propiedades de los cultivos, por medio de sensores que se encargan de medir la heterogeneidad de un campo (Fundación para la Innovación Agraria, 2008). De este modo, aplicar técnicas apropiadas de manejo de la información, se ha convertido en una gran necesidad en la industria agrícola teniendo como objetivo obtener una mayor eficiencia operativa.

La modernización y utilización de nuevas tecnologías para obtener datos de interés como sistemas de posicionamiento global (GPS), redes inalámbricas de sensores y drones, ofrecen un uso preciso y sostenible de insumos, fertilizantes y cantidad de agua que necesite un cultivo, siendo esto conocido como agricultura de precisión (Wang, Zhang, & Wang, 2006). Esta tecnificación puede también llevarse a cabo por medio del Internet de las Cosas (IoT). En IoT las cosas, es decir, dispositivos embebidos puedan estar idealmente disponibles en cualquier momento y en cualquier lugar a cualquier persona. En un sentido más técnico, IoT consiste en la integración de sensores y dispositivos en objetos cotidianos para que estos queden conectados a Internet a través de redes fijas e inalámbricas (Fundación de la Innovación Bankinter, 2011).

Una de las áreas de mayor aplicación del IoT es la meteorología, al usar sistemas con múltiples sensores para la monitorización en tiempo real de las condiciones climáticas, buscando obtener predicciones concisas y precisas. A pesar de los grandes avances agrícolas y tecnológicos de los últimos años, las condiciones atmosféricas, el impacto del cambio climático y fenómenos atmosféricos como el niño y la niña que generan desde fuertes sequías hasta incontrolables inundaciones, siguen siendo un gran problema que impacta a la industria generando altos riesgos y altos costos en todos los ámbitos, siendo necesario una adecuada previsión meteorológica con el fin de eliminar la necesidad de usar la intuición en la agricultura (Fundación de la Innovación Bankinter, 2011) (Lopez, Chavez, & Sánchez, 2017).

Las técnicas de minería de datos tienen como objetivo encontrar patrones en los datos que son valiosos e interesantes para determinar las posibles condiciones climáticas en los cultivos. Sin embargo, el problema para un agricultor que aún no ha implementado en sus cultivos algunas de estas técnicas de predicción y herramientas de agricultura de precisión, es el alto costo y la complejidad en interpretar la información obtenida al utilizar sistemas de posicionamiento global GPS o imágenes satelitales e incluso la utilización de drones. Por lo anterior, existe una gran demanda y necesidad de integrar un conjunto de componentes adecuados y de bajo costo organizados en una propuesta arquitectónica, dada la gran cantidad de tecnologías que vienen surgiendo para simplificar la obtención, almacenamiento y procesamiento de datos climatológicos, donde los factores más influyentes son la condición de sequedad y la alta temperatura por radiación solar. La demanda de predicciones climáticas concisas y precisas ha aumentado considerablemente en la industria agrícola, por lo que los agricultores desean saber tan pronto como sea posible, las condiciones atmosféricas en las que se encuentra su cultivo; además, de poder obtener predicciones de dichas condiciones e indicaciones de cuando es debido irrigar sus cultivos, para así reducir el uso inapropiado de recursos y mejorar el rendimiento de sus cultivos.

Teniendo en cuenta lo anterior, en este artículo se presenta como aporte la propuesta de una arquitectura para agricultura de precisión soportada en tecnologías abiertas para IoT y técnicas de minería de datos para la obtención, visualización y análisis en tiempo real de variables climáticas en un cultivo de prueba a pequeña escala. Esta propuesta es abordada desde tres vistas arquitectónicas asociadas al modelo de negocio, a la funcionalidad y a la implementación de la arquitectura para escenarios de agricultura de precisión. En estas vistas se ha tomado como referencia la estructura de capas de la arquitectura Lambda (Deshmane, 2015). De este modo, cada vista considera cuatro capas a saber: la de captura de datos, la de almacenamiento, la de procesamiento y la de consulta. La capa de captura de datos está formada por un conjunto de sensores cuyo objetivo es la obtención de variables climáticas asociadas al cultivo. La capa de almacenamiento se encarga de recopilar información en tiempo real a partir de las diferentes fuentes (sensores) distribuida en el cultivo. La capa de procesamiento tiene como función la predicción de información relevante a partir de los datos recopilados, usando en este caso técnicas de minería de datos y/o clasificación. Finalmente, la capa de consulta permite la visualización y seguimiento en tiempo real de las variables del cultivo y de las predicciones obtenidas en la capa de procesamiento.

Esta propuesta arquitectónica pretende servir de referencia para la implementación de servicio en el área de la agricultura, con el fin de mejorar la productividad agrícola, a través de la flexibilidad y funcionalidad que ofrece la combinación de herramientas de hardware y software libres para la implementación de servicios soportados en IoT en escenarios de agricultura de precisión. El resto del artículo está organizado de la siguiente forma: en la sección 2 se presentan un conjunto de trabajos relacionados que se consideraron en la presente investigación. En la sección 3 se presenta el marco conceptual con los conceptos más relevantes que se tuvieron en cuenta para el desarrollo de la presente investigación. En la sección 4 se describen cada una de las vistas de la propuesta arquitectónica; en la sección 5 se presenta la evaluación de la propuesta arquitectónica y finalmente en la sección 6 se presentan las conclusiones y trabajos futuros derivados de la presente investigación.

2. Trabajos Relacionados

En esta sección se presentan un conjunto de trabajos relacionados con la presente investigación, los cuales se enmarcan en la temática de arquitecturas IoT y sus aplicaciones en diferentes contextos que van desde el uso de este concepto en hogares inteligentes hasta el desarrollo de entornos de Cloud Computing para IoT.

En (Ungurean, Gaitan, & Gaitan, 2014) se presenta una arquitectura de IoT basada en las especificaciones de OPC.NET (Open Platform Communications) que puede usarse tanto en entornos industriales como en edificios inteligentes. El objetivo de los autores fue brindar una arquitectura formada por módulos principales: el servidor de datos y la aplicación de HMI (Human Machine Interface). El servidor obtiene datos de una red de sensores (Bus de campo) y envía comandos a los actuadores (tales como relés y electroválvulas) que están conectados en los buses de campo. La aplicación del servidor puede adquirir datos de redes inalámbricas basadas en ZigBee. Este trabajo aporta ideas importantes relacionadas con la manera de representar la información de interés al

usuario a través de distintos dispositivos móviles, así como la integración de sensores para la captura de datos de interés.

En (Datta, Bonnet, & Nikaein, 2014) se propone una arquitectura IoT centralizada para ofrecer nuevos servicios M2M (Machine to Machine), la cual permite la interacción en tiempo real entre clientes móviles y objetos inteligentes (sensores y actuadores) a través de una pasarela inalámbrica. Así, este artículo brinda las bases para realizar el procesamiento de los datos obtenidos.

En (Yashiro, Kobayashi, Koshizuka, & Sakamura, 2013) se propone la arquitectura uID-CoAP, una nueva arquitectura diseñada para alojar servicios IoT en sistemas embebidos comunes, como los aparatos de consumo habituales. Los autores combinan el protocolo de aplicación restringido (CoAP) con la arquitectura omnipresente (uID) que permite el reconocimiento de los datos requeridos para los servicios IoT. Mediante este artículo se corroboró la importancia de ofrecer aplicaciones prácticas soportadas en IoT sobre los sistemas embebidos existentes con poco esfuerzo computacional.

Múltiples proyectos se han desarrollado en torno a la interacción entre el IoT y la agricultura; donde se encuentra el proyecto desarrollado en la Universidad de Pekín, el cual diseña una arquitectura conformada por diferentes herramientas que generan la plataforma de infraestructura de información espacial, la plataforma de infraestructura de la IoT, la plataforma de gestión de la agricultura y el cliente móvil para un sistema de gestión de la agricultura de precisión PAMS; permitiendo a los usuarios supervisar la producción agrícola, donde el método de integración de módulo y software de código abierto ayudan a reducir el costo de desarrollo y mejorar la eficiencia del sistema (Li, 2012). Mediante este referente teórico se evidenció la aplicabilidad de tecnologías IoT para soportar el área de la agricultura.

Así, a partir de la literatura encontrada se plantea como brecha investigativa la necesidad de integrar dichos componentes en una arquitectura propia para entornos de agricultura de precisión. De este modo, en este artículo se hace una propuesta arquitectónica para IoT basada en el uso de plataformas de hardware libre, servicios web, bases de datos en línea o plataformas en la nube y protocolos de comunicación, implementándose en un entorno de agricultura de precisión el cual es un cultivo de vegetales a pequeña escala.

3. Marco Conceptual

En esta sección se presentan un conjunto de conceptos y tecnologías que se tuvieron en cuenta para el desarrollo del presente trabajo, dentro de los cuales se encuentran: agricultura de precisión, IoT, arquitectura Lambda, Framework Spark de Java y la aplicación de minería de datos Weka.

3.1. Agricultura de precisión

Tradicionalmente, la agricultura se practica realizando una tarea particular, tal como siembra o cosecha, en función de un horario predeterminado. Sin embargo, al recolectar datos en tiempo real sobre el clima, la calidad del suelo y del aire, la madurez de los cultivos e incluso los costos, la disponibilidad de los equipos y la mano de obra, puede hacerse un análisis predictivo para tomar decisiones más inteligentes, lo cual es conocido como agricultura de precisión (Ibañez, 2009). La implementación de técnicas de agricultura

de precisión se realiza con el objetivo de poder determinar con exactitud las necesidades de riego y de fertilizantes, las fases de desarrollo y de maduración de los productos, los puntos óptimos de siembra y de recolección, etc. Es decir, predecir adecuadamente las distintas etapas de producción de los cultivos. Para ello, es importante obtener toda la información posible del agua, del suelo, de las plantas y del ambiente.

3.2. IoT en la agricultura de precisión

El Internet de las Cosas (IoT) es una red de objetos tales como sensores y actuadores conectados a Internet a través de redes fijas e inalámbricas donde pueden capturar datos de forma autónoma y auto-configurable de forma inteligente basada en hechos del mundo físico; permitiendo que estos sistemas se conviertan en participantes activos en diversos procesos públicos, comerciales, científicos y personales (Fundación de la Innovación Bankinter, 2011).

El IoT es impulsado por una combinación de métodos y procesos que comprenden la utilización de sensores, una red y un dispositivo final, los cuales tienen como objetivo principal proporcionar al usuario final una representación de los datos de una manera visual, legible y práctica. Cuando se habla de IoT enfocado en la agricultura no sólo se hace referencia a dispositivos móviles como tabletas o drones, pues este concepto implica una interconexión de sistemas agrícolas de comunicaciones compatibles y colaborativos que comparten información (sensores, cámaras, lámparas, estaciones meteorológicas, etc.) (Fundación de la Innovación Bankinter, 2011).

3.3. Arquitectura Lambda

Una arquitectura de referencia estándar que comprende al IoT es conocida como la arquitectura Lambda (Deshmane, 2015), la cual consta de cuatro componentes a saber: Captura de Datos, capa de Velocidad o Procesamiento, la Capa de Almacenamiento o Seguridad y la Capa de Consulta. La capa de Captura de Datos, encargada del ingreso de la información, procesamiento de streaming o capa de velocidad, satisface todas las solicitudes que están sujetas a requisitos de baja latencia usando algoritmos rápidos e incrementales, considerando únicamente datos recientes. La capa de almacenamiento o lotes soporta la gestión de todo el conjunto de datos obtenidos, haciendo un histórico con estos. Finalmente, la capa de consulta donde se representa la información ya sea la obtenida en tiempo real o la parte almacenada.

La arquitectura Lambda intenta equilibrar la latencia, el rendimiento y la tolerancia a fallos mediante el procesamiento por lotes para proporcionar vistas completas y precisas de los datos por lotes, mientras que simultáneamente se utiliza el procesamiento de flujo en tiempo real para proporcionar vistas de datos en línea.

3.4. SparkJava Framework

Spark es un framework libre y de código abierto escrito en lenguaje Java para el desarrollo de aplicaciones web y dominios específicos. Es una alternativa a otros entornos de aplicaciones web en Java, como JAX-RS, framework Play y Spring MVC (Model-View-Controller). Se ejecuta en un servidor web embebido Jetty por defecto, pero se puede configurar para ejecutarse en otros servidores web. Spark fue inspirado en Sinatra, una librería de código abierto para aplicaciones web en lenguaje dominio- específico; no se

sigue el patrón modelo-vista-controlador utilizado en otros framework, como Spring MVC. En su lugar, Spark está destinado a la creación rápida de aplicaciones web en Java con el mínimo esfuerzo (Jurado, 2013). El framework Spark fue usado como pasarela de integración y procesamiento de los datos obtenidos a partir de los sensores.

3.5. Weka

Weka es un software libre desarrollado por la Universidad de Waikato (Nueva Zelanda), el cual cuenta con una extensa colección de algoritmos de aprendizaje de máquina implementados en Java, los cuales pueden ser aplicados sobre datos ya sea mediante la interfaz gráfica o para embeberlos dentro de cualquier aplicación. Además, Weka contiene las herramientas necesarias para realizar transformaciones sobre los datos, tareas de clasificación, regresión, clustering, asociación y visualización. Weka está diseñado como una herramienta orientada a la extensibilidad por lo que añadir nuevas funcionalidades es una tarea sencilla (Kulkarni & Kulkarni, 2016).

La licencia de Weka es GNU-GPL, lo que significa que este programa es de libre distribución. Además, ya que esta herramienta está desarrollada en Java, es independiente de la arquitectura, ya que funciona en cualquier plataforma sobre la que haya una máquina virtual Java disponible, permitiendo acceder a la biblioteca de sus propios programas Java o escribir y probar los esquemas de aprendizajes propios. En este trabajo se hizo uso de Weka para la implementación de un clasificador bayesiano, cuyo propósito es predecir la categoría de las instancias en función de una serie de atributos de entrada. En este caso, la clase es uno de los atributos simbólicos disponibles, el cual puede estar asociado a la condición climática o de riego del cultivo. El resultado de aplicar el algoritmo de clasificación se efectúa comparando la clase predicha con la clase real de las instancias.

4. Arquitectura Soportada en IoT

En esta sección se presenta la descripción de la arquitectura para agricultura de precisión soportada en IoT, la cual se representa por medio de tres vistas a saber: vista del negocio, vista funcional y vista de implementación (ver Figura 1). En la vista del negocio se presenta el

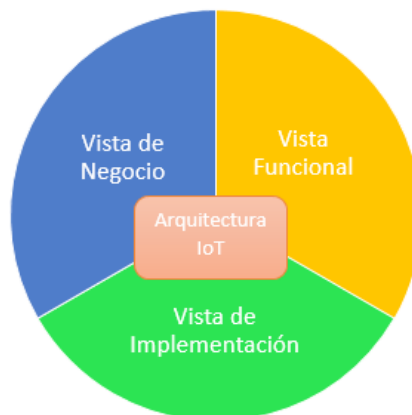


Figura 1 – Vistas de la arquitectura

modelo de negocio de la arquitectura en el contexto de la agricultura de precisión. En la vista funcional se describen los diferentes componentes funcionales de la propuesta arquitectónica y finalmente en la vista de implementación se presentan los diferentes componentes hardware y software escogidos para la construcción de la arquitectura. A continuación, se presenta la descripción de cada una de las capas de la propuesta arquitectónica.

4.1. Vista de Negocio

En esta vista de negocio se presenta la arquitectura IoT propuesta desde las cuatro capas asociadas a la arquitectura Lambda: capa de captura, capa de almacenamiento, capa de procesamiento y la capa de consulta. Como se muestra en la Figura 2, en la capa de captura se encuentran los sensores y módulos de adquisición de datos, así como los sistemas embebidos, microcontroladores o estaciones meteorológicas para la toma de las condiciones climáticas en el cultivo. La capa de almacenamiento está constituida por plataformas, servidores y base de datos que permiten guardar todos los datos obtenidos en la capa de captura. En la capa de procesamiento se extraen estos datos almacenados para realizar su respectivo análisis y obtener información de interés que puede ser aplicada en los cultivos. El componente final en la arquitectura es la capa de consulta, la cual le permite al agricultor visualizar en tiempo real el comportamiento de las variables asociadas al cultivo, así como las predicciones y recomendaciones a tomar de acuerdo a las necesidades del cultivo.



Figura 2 – Vista del negocio para la arquitectura IoT

4.2. Vista funcional

En esta sección se presenta la estructura funcional concebida para la propuesta arquitectónica fundamentada en la arquitectura Lambda. Así en la vista funcional se distinguen las cuatro capas Lambda: captura, almacenamiento, procesamiento y

consulta, ver Figura 3. La capa de captura de datos se encarga de realizar la adquisición de las variables meteorológicas a través de los sensores y el sistema embebido. La capa de almacenamiento tiene por función recibir los datos y almacenarlos en la plataforma en línea. La capa de procesamiento se encarga de analizar estos datos por medio de la herramienta de minería de datos. A través de la cual es posible la generación de predicciones y recomendaciones. Finalmente, la capa de consulta es la encargada de la monitorización de las condiciones del cultivo y visualización de la información entregada por la capa de procesamiento.

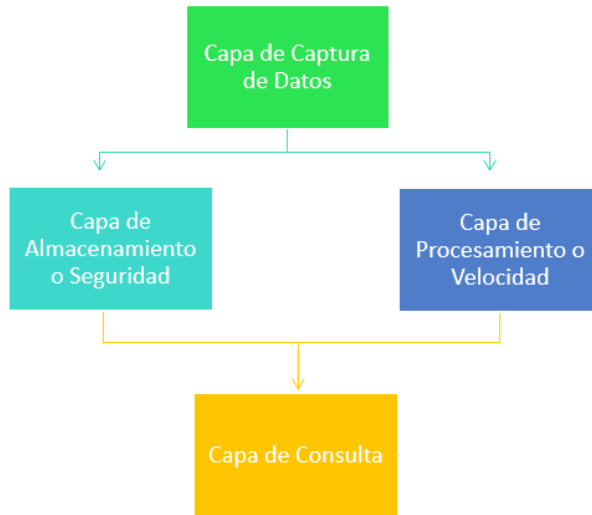


Figura 3 – Diagrama funcional del sistema IoT

En la figura 3, se presentan los diferentes módulos funcionales asociados a las diferentes capas de la propuesta arquitectónica. En la capa de captura, el flujo de datos entrantes se toma con el sistema de Captura de Datos (Sensores-Microcontrolador) y se puede enviar en dos direcciones, ya sea la capa de almacenamiento (base de datos en Línea, nube, servidores) o a la capa de procesamiento (herramientas Machine Learning). Esto garantiza que cualquier pérdida de datos se puede reproducir para garantizar que todos ellos se procesan al menos una vez. Esta información puede ser consultada a través de internet por medio de los dispositivos finales (servicios web, aplicaciones móviles) por medio de la capa de consulta.

La capa de almacenamiento o de seguridad brinda un archivo histórico, guardando todos los datos que se hayan recogido. Su almacenamiento podría ser para siempre, o se tendría que evaluar exactamente cuánto tiempo se desea mantener los datos (períodos de tiempo de meses o años). El almacenamiento para siempre soportaría análisis avanzados y los modelos predictivos que se desean implementar tendrían mayor exactitud para crear señales en la plataforma de transmisión y para consultas generales.

La capa de velocidad o procesamiento se define como una combinación de almacenes de colas, streaming y datos operativos. En la Arquitectura Lambda, la capa de velocidad es similar a la capa de almacenamiento a la hora de estimaciones analíticas, salvo que el proceso se realice en tiempo real. La capa de almacenamiento basa sus estimaciones en datos almacenados previamente, por lo que la responsabilidad de la capa de velocidad es obtener análisis en tiempo real a partir de datos de movimiento rápido.

Estas funciones descritas para cada una de las capas de la arquitectura propuesta, se pueden desglosar en sub-módulos funcionales como se presenta en el diagrama de flujo de la Figura 4, en el cual se representa la secuencia funcional de la arquitectura soportada en IoT para agricultura de precisión, de acuerdo a las capas provistas por la arquitectura Lambda.

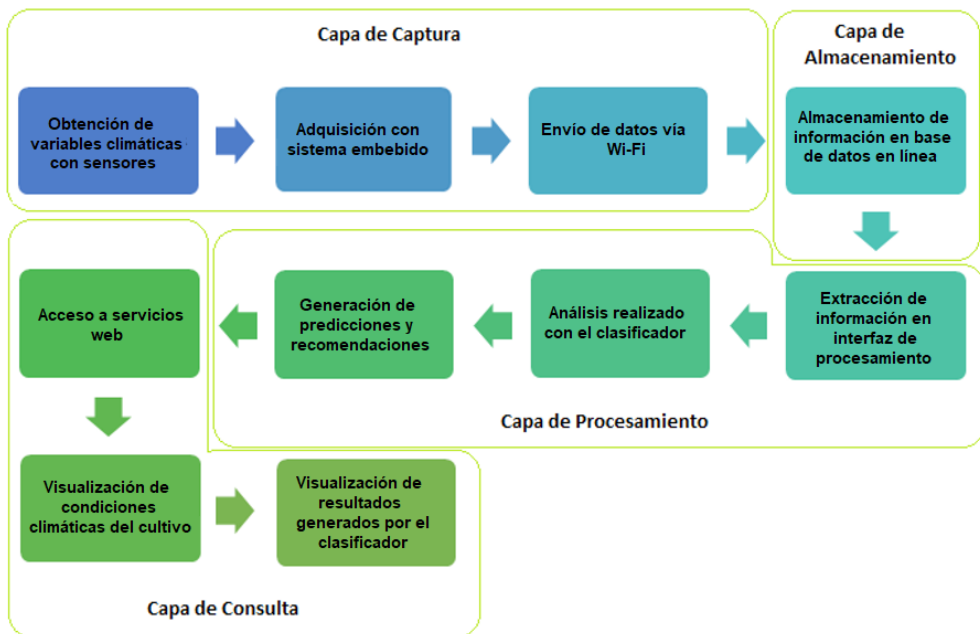


Figura 4 – Secuencia funcional de la arquitectura IoT

4.3. Vista de implementación

En esta sección se presenta detalladamente los componentes hardware y software escogidos para la implementación de la propuesta arquitectónica soportada en IoT y descrita en el presente artículo, ver Figura 5, los cuales están organizados de acuerdo a las capas de la arquitectura Lambda.

De acuerdo a la figura 5, en "1", se ubican sensores de temperatura, humedad relativa, humedad de suelo y luminosidad en el cultivo de prueba para realizar la adquisición de las variables climáticas correspondientes. En "2" por medio de la tarjeta de IoT Arduino YÚN, se configuran estos sensores con sus parámetros necesarios y se

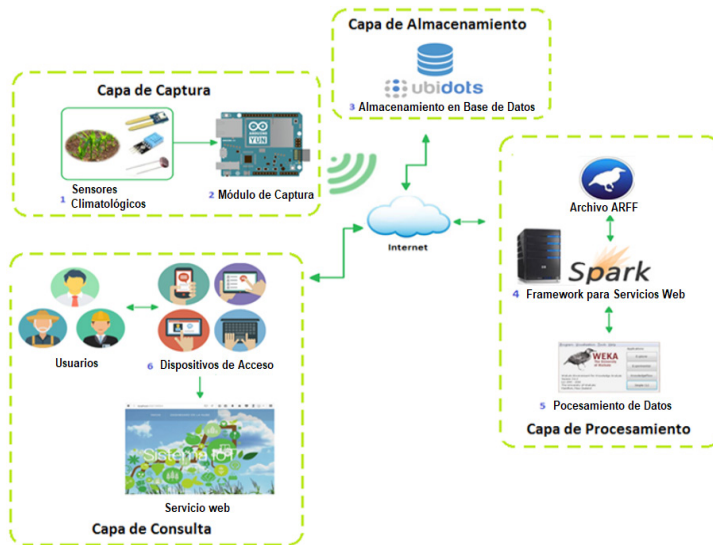


Figura 5 – Vista de implementación de la arquitectura IoT

establece la comunicación inalámbrica formando así la capa de captura de datos. A través de la tarjeta de captura se envía en tiempo real el registro de cada una de las variables asociadas a los sensores para ser almacenadas en “3”, por medio de la base de datos en línea Ubidots.

En “4” haciendo uso del framework de desarrollo web Spark, se implementa en Java un servicio web que permite realizar en línea el procesamiento de los datos capturados a partir de las anteriores capas. Para realizar el procesamiento de estos datos, al framework Spark integra en “5” la API provista por la herramienta de minería de datos Weka. Mediante esta API, se implementa un clasificador basado en el algoritmo de Naive Bayes con la información contenida en el archivo en formato arff (attribute-relation file formate), obteniendo así la capa de procesamiento y brindando las predicciones de las condiciones atmosféricas a tener en cuenta en el cultivo.

Finalmente, por medio del servicio web siendo visualizado en “6” por los usuarios, se genera la capa de consulta donde se monitorean las condiciones del cultivo y se proporcionan los resultados de predicciones entregados por el clasificador y las recomendaciones de interés.

A continuación, se presenta la descripción en detalle de cada uno de los componentes hardware y software asociados a las capas de la propuesta arquitectónica.

Capa de captura

En la figura 6 se presenta el esquema de conexión establecido para operar los sensores de temperatura, humedad relativa, humedad de suelo y luminosidad conectados junto a la tarjeta Arduino YÚN para la captura de las variables climáticas.

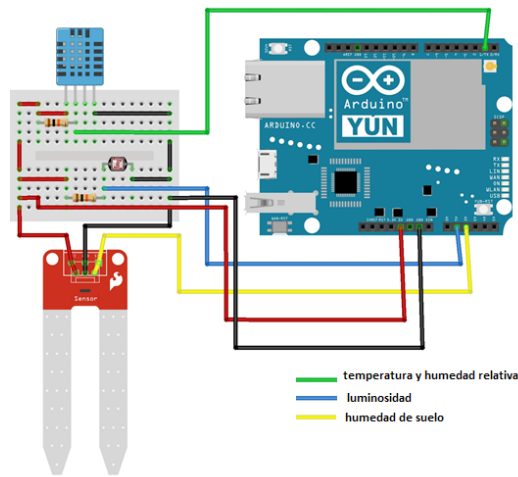


Figura 6 – Esquemas de conexión de sensores con microcontrolador

Para esta investigación se utiliza la tarjeta de IoT Arduino YÚN como el sistema embebido adecuado en la capa de captura de datos y para utilizar en el diseño de dispositivos conectados y en proyectos de IoT. Combina la potencia de Linux con la facilidad de uso de Arduino ya que al poseer comunicación y configuración vía Wi-Fi, permite enviar los datos de variables climáticas de forma inalámbrica a las otras capas de la arquitectura propuesta.

Los sensores utilizados son el DHT11, SEN9235P y un LDR, elegidos para esta investigación por su bajo costo cumpliendo con el requisito de componentes económicos para la implementación de la arquitectura para la agricultura de precisión. También se consideró la compatibilidad de estos componentes con el módulo Arduino YÚN y sus características de trabajo para operar en un cultivo controlado.

Capa de almacenamiento

Para establecer la propuesta arquitectónica IoT basada en Lambda, se debe contar con un gestor de almacenamiento que funcione como la capa de almacenamiento. Para lo anterior, se seleccionó la base de datos en línea Ubidots la cual es una herramienta especializada para aplicaciones IoT. Ofrece una plataforma para desarrolladores que permite capturar fácilmente datos de sensores para convertirlos en información útil, la cual brinda una adecuada compatibilidad con el módulo de captura Arduino YÚN y en su versión de prueba permite el almacenamiento de hasta 500.000 muestras.

En esta sección se presenta el módulo de administración de la base de datos en línea Ubidots, en el cual se presentan las diferentes variables climáticas capturadas a través de la tarjeta Arduino YÚN, ver Figura 7.

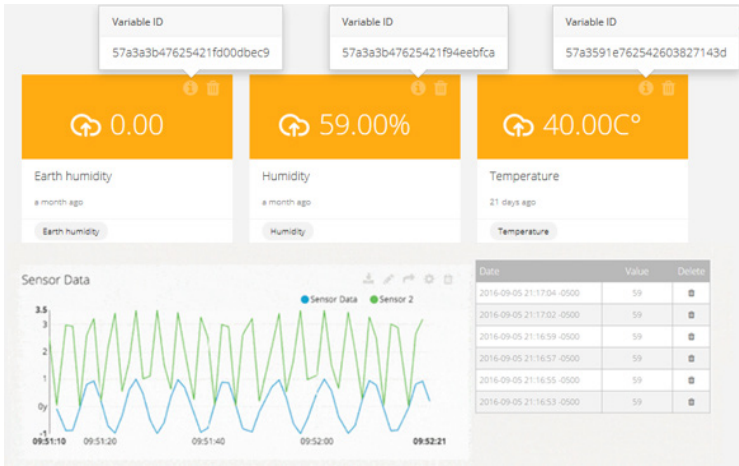


Figura 7 – Variables climáticas asociadas en base de datos en línea y su almacenamiento

Capa de Procesamiento

Por medio de la capa de procesamiento, los datos obtenidos de las variables climáticas pasan por un proceso de análisis mediante el algoritmo de aprendizaje bayesiano ofrecido por el software de minería de datos. Este algoritmo permite obtener las predicciones de las condiciones climáticas en el cultivo, así como recomendaciones de cuando es debido irrigar o no el cultivo, facilitando la toma de decisiones al agricultor y aportando a la reducción de insumos requeridos por el cultivo. Esta capa se comprende de dos etapas de procesamiento: la generación de un archivo en el formato específico de Weka y el procesamiento del archivo generado haciendo uso del software de minería de datos Weka.

Generación de archivo ARFF

Para procesar las variables climáticas con la herramienta de minería de datos Weka, se deben organizar la información de una manera específica, esto se realiza mediante un archivo ARFF, el cual es un formato de texto en ASCII que describe una lista de instancias compartiendo un juego de atributos; siendo este formato el único compatible para usarse en el software.

Estos archivos tienen dos secciones distintas, ver Figura 8: La primera sección es la información de cabecera la cual contiene el nombre de la relación, una lista de los atributos donde cada fila corresponde a una variable climática con los respectivos valores que este atributo puede tener, finalmente en esta sección se ubica la clase, en este caso asociada a la condición de riego del cultivo, siendo esta clase la que considera el software Weka para realizar las predicciones.

En la segunda sección del archivo ARFF, se presentan los datos organizados como instancias, donde cada columna corresponde al valor de la variable climática siguiendo el orden en que se ubicaron los atributos.

```

@relation meteorologia

@attribute humedadrelativa numeric
@attribute temperatura {baja, normal, alta}
@attribute humedadsuelo {seco, medio, humedo}
@attribute iluminacion {nublado, soleado, noche}
@attribute Class {si,no}

@data
59,alta,seco,soleado,si
59,alta,seco,soleado,si
59,alta,seco,soleado,si
59,normal,seco,nublado,si
59,baja,medio,soleado,no
59,alta,seco,noche,si
59,alta,medio,nublado,si
59,alta,medio,nublado,si
59,alta,humedo,nublado,no
59,alta,humedo,nublado,no
62,alta,humedo,soleado,no
62,alta,medio,soleado,si
62,alta,medio,soleado,si
62,alta,medio,nublado,si
62,alta,humedo,nublado,no

```

Figura 8 – Archivo ARFF generado para procesamiento

Análisis con software Weka

Una vez generado el archivo ARFF, este es analizado a través de la API provista por la herramienta Weka, considerando la secuencia presentada en la figura 9. La primera etapa en la utilización de esta herramienta es el pre-procesamiento, el cual consiste en la manipulación de vectores con contenido discreto de los valores sensados, donde al cargar el archivo con los datos necesarios, se identifica la clase a predecir y se pueden realizar filtrados de atributos, instancias o discretización. La segunda etapa consiste en la generación del clasificador utilizando un algoritmo de aprendizaje al cual se le entregan un conjunto de datos los cuales corresponden al contenido del archivo ARFF. Para esta investigación se hizo uso del algoritmo de Naive Bayes y los datos entregados para

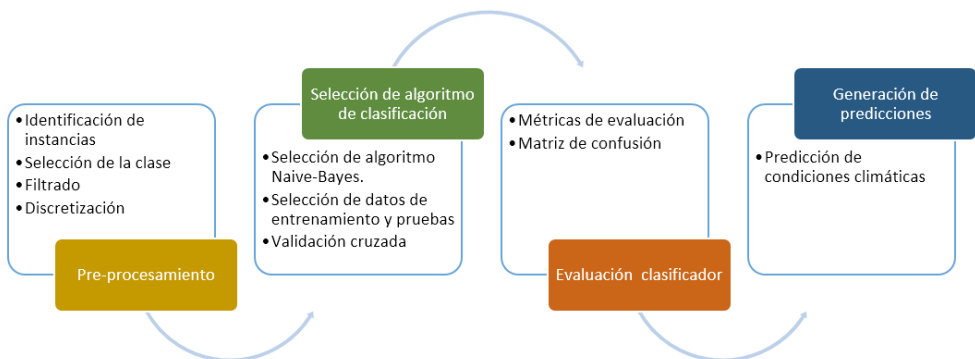


Figura 9 – Proceso de operación de software de minería de datos Weka

generar el clasificador se seleccionaron como conjunto de datos de entrenamiento y con validación cruzada.

En la tercera etapa se realiza la evaluación de este clasificador, lo cual puede definirse como el grado de concordancia entre los valores de la clase asignados por el clasificador y sus ubicaciones correctas según los datos entregados como referencia en el archivo ARFF. Para esto se presentan las métricas o información de ejecución entregada por Weka que permiten evaluar el rendimiento del modelo basado en el algoritmo de Naive Bayes. En la última etapa se representan las predicciones realizadas por el clasificador de la condición climática seleccionada.

Al finalizar este proceso, el software Weka entrega los resultados con las predicciones y recomendaciones obtenidas a partir de las condiciones climáticas del cultivo, las cuales son visualizados en la siguiente capa de la arquitectura propuesta.

Capa de consulta

Para completar la etapa final de la arquitectura, se desarrolló una interfaz gráfica de acceso al servicio web de procesamiento provisto por el framework Spark. La interfaz desarrollada permite visualizar en tiempo real en forma de tabla y de manera gráfica, los datos climáticos de temperatura, humedad relativa, humedad de suelo e iluminación que está tomando el componente de captura de datos, ver Fig. 10. Además, esta interfaz posibilita la consulta en línea del histórico con los datos almacenados en la base de datos Ubidots, así como las predicciones y recomendaciones para que el agricultor puede tomar decisión acerca del riego del cultivo.

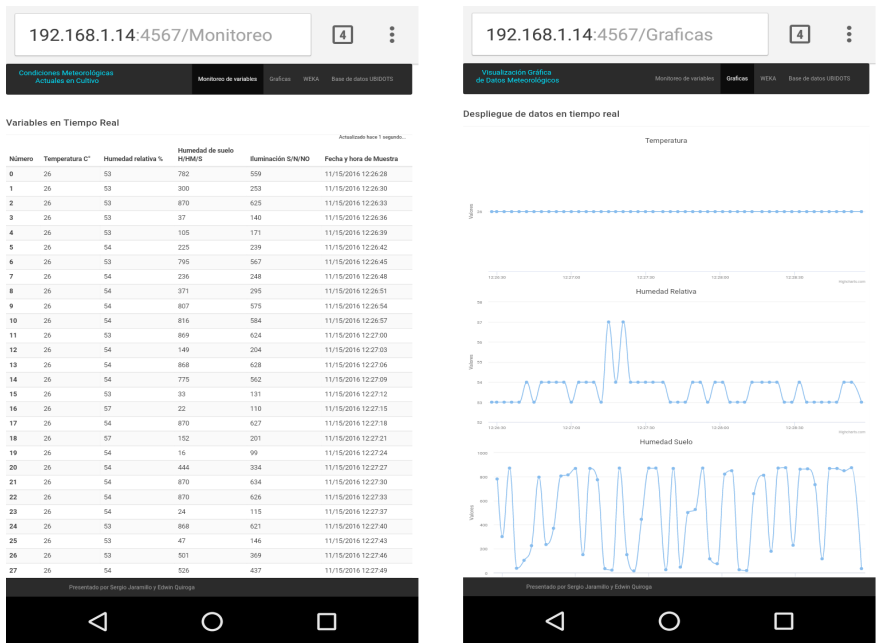


Figura 10 – Visualización de pestañas del servicio web para la consulta de la información

5. Evaluación de la propuesta arquitectónica

Finalmente, a modo de evaluación de la propuesta arquitectónica, se evaluó el componente principal de la arquitectura, el cual corresponde al módulo de procesamiento. Este componente fue sometido a pruebas de carga con el fin de determinar su capacidad y su tiempo de respuesta. Estas pruebas realizadas permiten obtener los límites bajo los cuales el módulo de procesamiento se comportará de manera apropiada y permite saber cuándo se necesitan implementar diferentes estrategias de optimización ya sea a niveles de configuración del servidor o de balanceo de carga.

Así, en las gráficas de las figuras 11 y 12 muestran los tiempos de respuesta para las pruebas de estrés realizadas sobre el módulo de procesamiento, las cuales se llevaron a cabo mediante la variación incremental de 100 a 1000 peticiones concurrentes. En la figura 12 se presenta la prueba para la URL Monitoreo, en la cual se realiza la extracción de la información de la base de datos Ubidots para realizar la tabulación de los datos climáticos en tiempo real. En la figura 13 se presentan los tiempos de respuesta para la prueba de estrés realizada a la URL WEKA, provista para la generación y visualización de las predicciones.

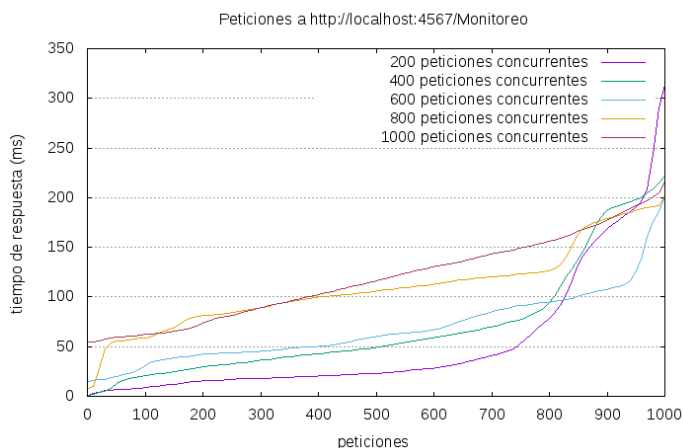


Figura 11 – Tiempos de respuesta para prueba de estrés en capa de procesamiento

En ambas pruebas de carga se corrobora como los tiempos de respuesta son más altos a medida que aumentan las peticiones concurrentes. En el caso de la URL Monitoreo en la figura 11, se tiene un tiempo de 2ms para 20 peticiones concurrentes y este tiempo aumenta hasta 70 ms cuando se tienen un número de 100 peticiones concurrentes. A medida que aumentan las peticiones concurrentes de 800 a 1000 los tiempos de respuesta aumentan de 100ms a 300ms. Esto es, a medida que la aplicación presenta tráfico continuo con un número de clientes en aumento, tiende a darse un retardo al sistema, afectando la carga, la actualización y la visualización de datos en tiempo real.

Los tiempos de respuesta de la URL WEKA en la figura 12, están entre 200ms para 200 peticiones concurrentes y hasta 3 segundos para 1000 peticiones concurrentes.

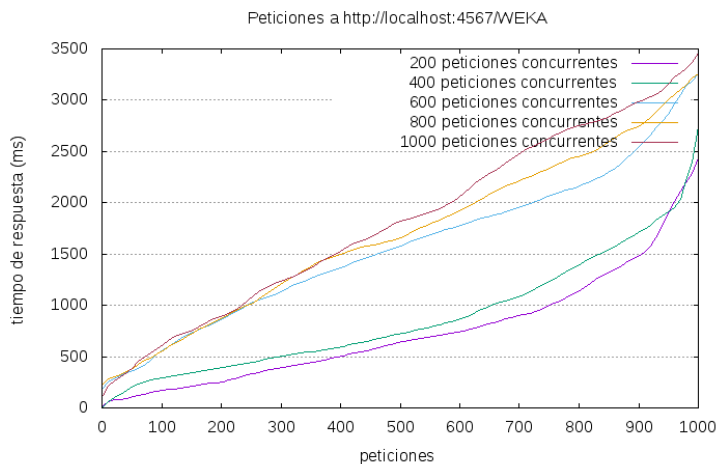


Figura 12 – Pruebas de estrés aplicadas al servicio web en la capa de consulta

Estos tiempos obtenidos se deben al procesamiento realizado desde la interfaz gráfica la cual realiza la extracción de la información desde Ubidots, la discretización de las variables meteorológicas, el pre-procesamiento, la clasificación y la evaluación.

Al observar el análisis de las pruebas de estrés efectuadas al módulo de procesamiento, el tiempo de respuesta máximo obtenido efectuando 1000 peticiones concurrentes es de 3000ms, el cual es un valor que no afecta la operación del sistema en el entorno de agricultura de precisión puesto que este no se somete a un alto número de usuarios accediendo simultáneamente. Además, el tiempo de respuesta máximo para una cantidad inferior de usuarios es de 50 ms el cual es aceptable para ofrecer toda la información solicitada en el momento deseado.

6. Conclusiones y trabajos futuros

Aprovechando la gran flexibilidad y funcionalidad que ofrece la combinación de herramientas hardware de bajo costo y software libre, se ha conseguido el diseño, implementación y despliegue de una arquitectura soportada en IoT, la cual se aborda desde tres vistas arquitectónicas de modelo de negocio, funcionalidad e implementación en un escenario de agricultura de precisión. Además, se despliegan cuatro capas de operación comprendiendo la captura de las variables climáticas, el almacenamiento de esta información, su procesamiento y su consulta; sirviendo como referencia para los agricultores y la comunidad interesada en utilizar nuevas tecnológicas, técnicas y servicios basados en IoT que buscar mejorar la productividad agrícola.

Se ha logrado integrar todos los componentes propuestos en esta investigación en torno a una propuesta arquitectónica, que parte de la arquitectura Lambda, teniendo el módulo de captura Arduino YÚN como componente de adquisición de datos, la base de datos en línea Ubidots correspondiente a la etapa de almacenamiento, la herramienta de minería

de datos Weka utilizada como la etapa de procesamiento y el framework de diseño web Spark de Java para la etapa de consulta; dando como resultado una arquitectura soportada en herramientas IoT.

Como trabajo futuro, se pretende evaluar esta propuesta arquitectónica en otros contextos de aplicación donde se requiera la captura de altos volúmenes de datos para realizar un procesamiento de estos en tiempo real, de tal modo que se consideren contextos tales como la educación, el marketing, ciudades inteligentes y entornos de desarrollo IoT.

Referencias

- Datta, S. K., Bonnet, C., & Nikaein, N. (2014). An IoT gateway centric architecture to provide novel M2M services. In *Proceedings of the 2014 IEEE World Forum on Internet of Things (WF-IoT)*, (pp. 514–519). Seoul. doi:10.1109/WF-IoT.2014.6803221
- Deshmane, S. (2015). *Using the Lambda Architecture on a Big Data Platform to Improve Mobile Campaign Management*. Retrieved from <http://www.talentica.com/pdf/Big-Data-Using-Lambda-Architecture.pdf>.
- Fundación de la Innovación Bankinter. (2011). *El Internet de las Cosas - En un mundo conectado de objetos inteligentes*. Retrieved from http://www.belt.es/expertos/imagenes/XV_FTF_El_internet_de_las_cosas.pdf
- Fundación para la Innovación Agraria. (2008). *Tecnologías aplicables en agricultura de precisión*. Retrieved from http://wapp.corfo.cl/ticrural/descargas/I+D+I_3_TECNOLOGIAS_AGRICULTURA_PRECISION.pdf
- Ibañes, J. (2009, Marzo 9). Agricultura de Precisión y Suelos. Retrieved from <http://www.madrimasd.org/blogs/universo/2009/03/09/114118>
- Jurado, J. (2013, Marzo). Spark: Framework ligero en Java. Retrieved from <http://www.happyminds.es/spark-framework-ligero-en-java/#sthash.ki1BKXfK.dpbs>
- Kulkarni, E., & Kulkarni, R. B. (2016). WEKA Powerful Tool in Data Mining. *International Journal of Computer Applications*, 9–15. Retrieved from <http://research.ijcaonline.org/rtdm2016/number2/rtdm2575.pdf>
- Li, S. (2012). Application of the Internet of Things Technology in Precision Agriculture Irrigation Systems. In *Proceedings of the 2012 International Conference on Computer Science and Service System* (pp. 1009–1013). Nanjing: IEEE. doi:10.1109/CSSS.2012.256
- Lopez, J., Chavez, J., & Sánchez, A. (2017). Modelado de una red de sensores y actuadores inalámbrica para aplicaciones en agricultura de precisión. In *Proceedings of the Humanitarian Technology Conference (MHTC), IEEE Mexican*. Puebla: IEEE.
- Yashiro, T., Kobayashi, S., Koshizuka, N., & Sakamura, K. (2013). An Internet of Things (IoT) architecture for embedded appliances. In *Proceedings of the Humanitarian Technology Conference (R10-HTC), 2013 IEEE Region 10* (pp. 314–319). Sendai: IEEE. doi:10.1109/R10-HTC.2013.6669062

- Ungurean, I., Gaitan, N.-C., & Gaitan, V. (2014). An IoT architecture for things from industrial environment. In *Proceedings of the Communications (COMM), 2014 10th International Conference on*. Bucharest - Romania: IEEE. doi:10.1109/ICComm.2014.6866713
- Wang, N., Zhang, N., & Wang, M. (2006). Wireless sensors in agriculture and food industry—Recent development and future. *Computers and Electronics in Agriculture*, 50(1), 1–14.

Las redes sociales web como herramienta para gestionar información en procesos de co-creación

Liliana González-Palacio ¹, Mauricio González-Palacio ², Jairo Ortiz-Pabón¹

ligonzalez@udem.edu.co, magonzalez@udem.edu.co, hortiz@udem.edu.co

¹ Ingeniería de sistemas- Facultad de ingeniería- Universidad de Medellín, Carrera 87 Nro. 30-65, 50026, Medellín, Colombia

² Ingeniería en telecomunicaciones- Facultad de ingeniería- Universidad de Medellín, Carrera 87 Nro. 30-65, 50026, Medellín, Colombia

DOI: 10.17013/risti.24.57-68

Resumen: Los procesos de innovación abierta emplean la co-creación de valor para viabilizar la interacción y colaboración entre agentes internos y externos a la organización. El uso de las tecnologías de información y comunicación hace posible que participantes distribuidos geográficamente y con limitantes de tiempo puedan aportar sus ideas en cualquier momento y lugar. Particularmente las redes sociales surgen como una alternativa interesante en la co-creación, ya que facilitan el involucramiento de colectivos identificados con las mismas necesidades y problemáticas. También estimulan el intercambio dinámico de experiencias en contextos de variada complejidad. En esta investigación se propone un protocolo para seleccionar una red social acorde a escenarios específicos de aplicación de la co-creación de valor. La iniciativa es validada mediante un caso real donde se evidencia que el uso de una red social para una construcción colectiva favorece la sinergia entre participantes heterogéneos.

Palabras-clave: Redes sociales, innovación abierta, co-creación, colaboración, interacción.

Social networks as a tool to manage information in co-creation processes

Abstract: Open innovation processes employ the co-creation of value for the interaction and collaboration between internal and external agents to the organization. The use of information and communication technologies enables stakeholders distributed geographically and with time constraints to express their ideas anywhere and anytime. Particularly, social networks emerge as an interesting alternative in co-creation, due to they facilitate the involvement of groups identified with the same needs and problems. They also encourage the dynamic exchange of experience in contexts of different kinds of complexity. This research proposes a protocol to select a social network according to specific application scenarios of co-creation of value. The initiative is validated by a real case where is evident that the use of a social network for collective construction promotes synergy between heterogeneous stakeholders.

Keywords: social networks, open innovation, co-creation, collaboration, interaction.

1. Introducción

En el paradigma tradicional de innovación las compañías siguen su cadena de valor ubicando al cliente en un rol pasivo. Hoy, la innovación abierta estimula la capacidad innovadora, acelera la transferencia de conocimiento e intensifica la capacidad de colaboración endógena y exógena de las empresas (Chesbrough, 2006; Enkel, Gassmann, & Chesbrough, 2009). La innovación ahora es pensada bajo un esquema abierto (Chesbrough, 2006) donde actores y agentes externos co-crean valor en todo el proceso y son escuchados por las compañías que fabrican nuevas ofertas, las cuales a su vez reciben aportes continuos y retroalimentación.

La innovación abierta (IA) es un modelo mediante el cual las empresas hacen uso intensivo de conocimiento interno y externo, con el objetivo de potenciar y acelerar su innovación interna y expandir los mercados (Chesbrough, 2006). Esta dinámica para innovar requiere la incorporación de medios tecnológicos adecuados para la interacción (Enkel, et al., 2009). El concepto de red social ha evolucionado mostrándose como una herramienta poderosa para facilitar procesos de co-creación de valor. Se evidencian vacíos importantes al seleccionar la red social adecuada, y finalmente la elección se hace sin tener criterios claros.

En esta investigación se propone un protocolo para seleccionar una red social acorde a las necesidades específicas de un proceso de co-creación de valor. Para el efecto se inicia con el levantamiento de un referente conceptual y la revisión de literatura, que permiten analizar la problemática. Posteriormente se presenta una propuesta de solución a las carencias identificadas. La iniciativa es validada mediante un caso real.

El artículo se organiza como sigue: la sección 2 muestra el fundamento teórico que sustenta la propuesta. La sección 3 contiene el trabajo relacionado. La problemática es descrita en el apartado 4. La sección 5 detalla el protocolo para la selección de una red social apropiada acorde a un ejercicio de co-creación específico. La sección 6 describe la validación. Finalmente se presentan las conclusiones, trabajo futuro, agradecimientos y referencias.

2. Marco teórico

Esta investigación se enmarca en los conceptos de innovación, co-creación, y redes sociales. A continuación se explican brevemente (Gonzalez, González, Echeverri, & Urrego, 2013; González & González, 2015):

- Innovación: según el manual de Oslo se define como la implementación de un producto (bien, servicio, proceso) nuevo o significativamente mejorado que se incorpora en el mercado (OECD, 2005).
- Co-creación: hace alusión a la generación conjunta de valor a partir de la interacción entre agentes internos y externos a la organización. Esto se potencia mediante los principios de diálogo continuo, acceso a la información, riesgo/beneficio y transparencia (Prahalad & Ramaswamy, 2004). Este concepto

está íntimamente ligado con la materialización de la innovación abierta en las organizaciones. Las características enunciadas exigen un manejo adecuado de aspectos tecnológicos y de seguridad, al tiempo que requieren un alto nivel de colaboración y comunicación entre los participantes (Ormaetxe, 2008).

- Redes sociales: originalmente hacen referencia a una reunión de personas, conocidas o desconocidas, que interactuarán entre sí, redefiniendo al grupo y retroalimentándolo en lo que podría asimilarse a una dinámica de trabajo, o bien a una estructura auto-organizada de personas compartiendo, manipulando y transformando información colaborativamente (Raghavan, 2002). Las redes sociales pueden definirse, como asociaciones de personas ligadas por motivos heterogéneos y que conforman una estructura compuesta por nodos unidos entre ellos por más de un tipo de relación (Requena, 2008). Los usuarios frecuentes de Facebook, o de otra red de características similares, suelen argumentar que se trata de un medio eficiente para mantener o renovar vínculos con amigos, familiares, instituciones u otros grupos de interés; pero más allá de esa percepción, algunos investigadores han equiparado el alcance de esa red social con el de un complejo sistema abierto y en construcción permanente que involucra a colectivos identificados con las mismas necesidades y problemáticas, donde además se estimula el intercambio dinámico de experiencias, conocimientos y muchos otros aspectos de la vida en contextos de variada complejidad (Anria, 2009). En esa misma dirección, esta dinámica de interacción permite a grupos e individuos, a través del mismo modelo de organización y disposición de infinidad de informaciones allí contenidas, potenciar sus recursos cognitivos (Caldevilla, 2010; Maamar & Badr, 2009; Tavante & Coutinho, 2011). Las redes sociales se vislumbran como una poderosa herramienta para facilitar la co-creación.

3. Trabajo relacionado

Esta sección da cuenta de una exploración sobre el uso de las redes sociales web en procesos de innovación abierta y co-creación de valor, o sectores afines.

Los países cuyas empresas están más avanzadas en el uso de las herramientas Web con fines de innovación colectiva y como parte de su estrategia empresarial son Estados Unidos, Reino Unido, Alemania, China e India (Nafria, 2008). Mientras que el entretenimiento y el turismo son algunos de los sectores que más aprovechan las ventajas que ofrece este tipo de tecnología comunicacional.

Profundizando en algunas iniciativas, (Cohen & Clemens, 2005) proponen una red social web para facilitar la colaboración creativa. La idea fue validada en la empresa consultora Razorfish, dedicada servicios de fidelización de marca. Dentro de su estudio se involucraron cincuenta diseñadores de interfaces, arquitectos de información y en general, ingenieros de sistemas. Como propuesta se creó una red social web denominada “Peers” para promover internamente la solución al problema organizacional en el cual la información no está visible a todos los integrantes del equipo, debido a la carencia de interacciones entre los mismos. Dicha red tiene un formato de blog compartiendo espacios de trabajo, además incluye una forma para votar por los aportes de cada usuario (clasificándolos de acuerdo a su relevancia).

(Lim, Damian, & Finkelstein, 2011) diseñaron una red social para apoyar los procesos de captura de requisitos (centrados en los usuarios) en diseño de software. Lograron evidenciar una mejora en la comunicación entre todos los participantes del diseño de un nuevo proyecto, en términos de fluidez, efectividad y trazabilidad.

(Tomsic & Suthers, 2005) llevaron a cabo un estudio observando las herramientas de interacción comunes dentro del departamento de policía de Honolulu (teléfono, correo electrónico y cara a cara). Luego analizaron cómo la inclusión de una red social mejora la comunicación y encontraron un aumento significativo.

(Murphy, Sheth, Kaiser, & Wilcox, 2008) sostienen que las herramientas para trabajo colaborativo en labores científicas se basan en la capacidad de compartir información, mas no conocimiento explícito. Argumentan además que las redes sociales son un medio ideal para alcanzar una comunicación más efectiva entre los participantes. Diseñaron la herramienta GenSpace cuya estructura tipo red social incluye mensajería instantánea, sugerencias de posibles contactos y clasificación de perfiles (básicos y expertos).

(Althoff, Jindal, & Leskovec, 2017) identifican en el uso de las redes sociales WEB seis características fundamentales: (1) son dinámicas, puesto que los contenidos se actualizan de forma constante; (2) son colaborativas, pues se elaboran por un grupo de personas; son simples e intuitivas; pueden utilizarse sin necesidad de instalar nada en el ordenador, ya que la web es la plataforma; (5) poseen un entorno amigable e interactivo y los usuarios tienen capacidad de gestionar qué, cuándo y cómo publicar.

(Aiello & Ruffo, 2012) se concentran en la seguridad en redes sociales y destacan la necesidad de potenciarla este aspecto facilitando interacciones confiables entre participantes. Así diseñan un framework para ser incluido en una red social, destacando atributos tales como la confidencialidad, la privacidad de las contribuciones, entre otros.

El estudio de Yang y Chen se enfoca en demostrar que las redes sociales facilitan el intercambio de conocimientos posibilitando que los participantes encuentren contenido relevante y compartan alrededor de temas específicos (Yang & Chen, 2008). Los autores también descubren un incremento en la creatividad por tratarse de una actividad social en la cual se genera conocimiento colectivo mediante la interacción. Además, proponen una arquitectura P2P para configurar redes sociales web.

(Anria, 2009) relaciona las tecnologías más usadas en la interacción a través de redes sociales, además de categorizar la comunicación mediada por el computador en: Uno a uno (Ejemplo: Mensajería instantánea para interacción privada), uno a muchos (Ejemplo: páginas web o blogs) y muchos a muchos (Ejemplo: wikis o Whiteboards). Este autor también enuncia las características fundamentales de una red social web y por último enuncia los beneficios y riesgos asociados al uso de estas herramientas en un entorno empresarial.

Dall'Asta y otros estudian la interacción en redes sociales para hacer analogías con la teoría de juegos (Dall'Asta, Marsili, & Pin, 2012). Se concentran en las diferentes topologías de red y su conformación en términos de grafos, para asociar estas estructuras a características que facilitan la interacción.

La innovación es esencial para que las empresas sobrevivan en la economía global de hoy. La co-creación es una estrategia donde clientes y proveedores participan en el proceso de innovación. Los medios sociales son una herramienta perfecta para co-crear

valor ya que promueven la comunicación entre la organización y el cliente de una forma fluida e intuitiva (Halale, Gangadharan, & Uden, 2015).

4. Problemática

Las organizaciones están empleando cada vez más las redes sociales en los negocios. Este fenómeno también atrae la atención de los investigadores. Los medios de comunicación social se utilizan en la co-creación de valor para permitir a los consumidores tomar un papel activo y co-crear valor para las empresas. Sin embargo, la forma en que los medios sociales pueden mejorar la capacidad de co-creación de valor de una empresa no ha sido estudiada en la misma medida (Wang, Li, & Suomi, 2016).

Es evidente que las redes sociales web son herramientas usadas en diversos sectores productivos para facilitar la interacción de participantes que comparten intereses, gustos y se encuentran distribuidos geográficamente. Al explorar el dominio de la co-creación mediante redes sociales web no se evidenciaron propuestas para elegir de una manera sistemática la red social que mejor se ajusta a un escenario particular.

De allí surge la idea de proveer un protocolo de selección de una red social con fines de co-creación, empleando para el efecto un conjunto de criterios que hagan el ejercicio replicable y argumentable. Esto debería mejorar indicadores como la cantidad de ideas recopiladas durante la co-creación de valor.

5. Solución propuesta

Se propone el siguiente método (figura 1) para la selección de una red social durante un ejercicio de co-creación.

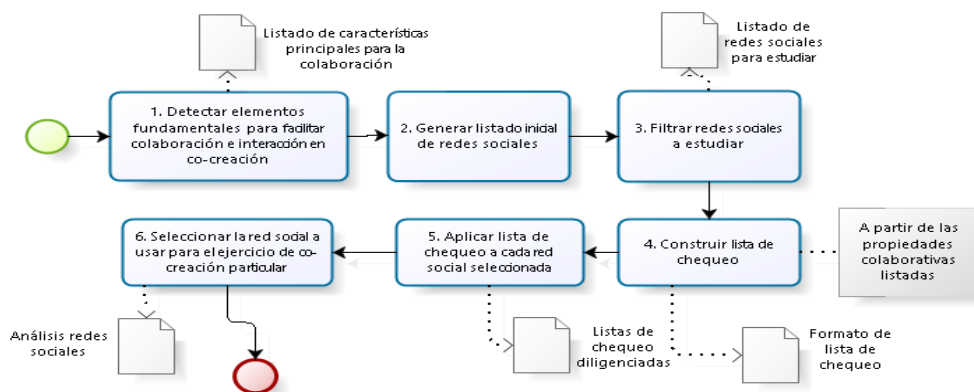


Figura 1 – Proceso para la selección de redes sociales en ejercicios de co-creación

Para la detección de las características de trabajo colaborativo, y sin perder de vista las particularidades de un ejercicio de co-creación, se sugiere incluir las siguientes características (González, González, & Urrego, 2013): organización y categorización de contactos (compartir, importar/exportar, conformación de grupos); sugerencias y recomendaciones de contactos (Cohen & Clemens, 2005; Lim, et al., 2011); generación de comentarios, y trazabilidad sobre éstos; intercambio de archivos de diferentes tipos (Murphy, et al., 2008); modificación en línea de los archivos compartidos (Aiello & Ruffo, 2012); gestión de versiones de los archivos modificados (Reinhardt, Moi, & Varlemann, 2009; Yang & Chen, 2008); notificaciones a los usuarios sobre cambios en los archivos; generación de blogs; programación de eventos (una vez, periódicos, rango de fechas); notificaciones y recordatorios de los eventos usando varios medios (correo, mensajes al móvil, ventanas emergentes, etc); generación de listas de distribución (Cohen & Clemens, 2005; Lim, et al., 2011); despliegue de publicidad; despliegue de calendario (con posibilidad de compartir con otros usuarios); realización de video-llamadas y webinar; intercambio de mensajes de forma síncrona y asíncrona; búsqueda de elementos (archivos, eventos, contactos); selección de estados (disponible, ocupado, ausente, desconectado, etc); esquema de calificación de usuarios de acuerdo a participación; mecanismo de premiación o bonificación para usuarios (Cohen & Clemens, 2005; Lim, et al., 2011); exportación e integración con otras redes sociales; publicación de contenidos (Reinhardt, et al., 2009; Yang & Chen, 2008); soporte para ambientes móviles; esquema de privacidad adecuado.

El siguiente paso es seleccionar el conjunto de redes sociales desde donde se elegirá una para hacer co-creación. Esta no es una labor trivial dada la cantidad exorbitante de redes sociales existentes. Para el efecto se sugiere revisar los listados y reportes ofrecidos por instituciones (en diversos países) como:

Institución	Descripción
Elogia IPSOFACTO (Elogia-Ipsufacto, 2015)	Son expertos en marketing para comercio electrónico. En conjunto con IAB Research Institute hacen estudios periódicos para evaluar cómo han evolucionado los principales indicadores de las redes sociales más relevantes.
ComScore (ComScore, 2015)	Se encargan de hacer investigación de marketing en Internet, ofreciendo servicios de información para muchas de las mayores empresas de Internet. Los análisis son hechos en diversos países, atendiendo el comportamiento variado de los usuarios.
EBizMBA (eBizMBA, 2015)	Rango generado a partir del Alexa Ranking Global Traffic y EE.UU. ranking

Tabla 1 – Reportes sobre popularidad en redes sociales web

Se pueden usar criterios como la frecuencia de uso de la red social, el tipo de interacción que facilita, los públicos que usan la herramienta a menudo, entre otros, para hacer la preselección de redes sociales a explorar con mayor detalle (Roch & Mosconi, 2016).

A este subconjunto de herramientas para la interacción se le debe aplicar el checklist que se indica en la tabla 2:

Categoría	Descripción
1	Espacio compartido de trabajo. Preguntas asociadas: ¿Es posible compartir documentos, compartir notas, compartir medios (videos, imágenes, presentaciones), generar blogs y wikis?
2	Calendario compartido. Preguntas asociadas: ¿Se puede tener una agenda común, generar eventos compartidos en grupos, realizar seguimiento a una tarea particular a través de notificaciones automáticas?
3	Mensajería y videoconferencia Preguntas asociadas: ¿Se cuenta con un sistema similar al correo electrónico?; ¿existe un concepto similar al de un foro?; ¿existe un chat online? ¿El chat permite videoconferencia?
4	Configuración de perfiles / privacidad. Preguntas asociadas: ¿El usuario puede configurar su perfil propio, unirse a un grupo de interés particular y/o restringir a otros el acceso a su perfil?
5	Recuperación de información. Preguntas asociadas: ¿Es posible generar etiquetas (o tags) de la información publicada por el usuario sobre el contenido compartido?; ¿Es posible organizar la información de acuerdo a una marca de hora y fecha?
6	Herramientas para medir la colaboración. Preguntas asociadas: ¿Existe un mecanismo explícito para aprobar o desaprobado un pensamiento, conducta, aporte o idea de otro usuario?; ¿El aporte puede ser re-enviado?
7	Interacción entre usuarios. Preguntas asociadas: ¿Se pueden compartir contactos? ¿se generan sugerencias de posibles contactos?

Tabla 2 – Lista de chequeo para la selección de redes sociales web

Se debe explorar la presencia o ausencia de cada categoría de la lista de chequeo anterior y de esta manera encontrar cuál de las redes sociales exploradas logra agrupar mayor cantidad de características favorables para un ejercicio de co-creación. De esta forma se determinará la herramienta a usar.

Siguiendo este método se tendrá un subconjunto de redes sociales con un ranking que indica cuál de ellas favorece más el trabajo colaborativo.

6. Validación

El escenario de validación fue un ejercicio de co-creación para aportar ideas sobre los elementos esenciales de un modelo de comercialización tecnológica en la región antioqueña (en Colombia). Los participantes fueron expertos en comercialización de tecnología de diversos países. Se requería elegir una herramienta adecuada para interacción virtual. Junto con los interesados en el ejercicio se definió que una red social web cubría los requerimientos para facilitar el intercambio de ideas. Para elegir cuál red social se usó el protocolo presentado en este artículo. Algunas características adicionales del escenario se indican a continuación:

- Se buscó proporcionar un medio de interacción entre expertos a escala global facilitando el espacio virtual para compartir experiencias y conocimiento informal entre los participantes.

- Los participantes estaban distribuidos geográficamente en países distantes.
- Se usó el inglés como idioma para la interacción.

Guiados por el método presentado anteriormente, en el primer paso se retomó el listado de características colaborativas sugerido sin encontrar otro aspecto adicional.

En el segundo paso se buscó un subconjunto de redes sociales representativas. Para ello se revisaron 12 reportes de las instituciones enunciadas en el método objeto de este artículo (Elogia IPSOFACTO en conjunto con IAB Research Institute, ComScore en países como Brasil, México, Argentina, Chile, Venezuela, Perú (Yanqui & Quispe, 2015), Colombia, EBizMBA). Se usó como criterio de selección las redes que tuvieran un porcentaje mayor de aparición en los reportes analizados, tal como se indica en la tabla 3:

#	Red social	% aparición	#	Red social	% aparición
1	Facebook	52.3%	6	MySpace	3.8%
2	Twitter	13.8%	7	Badoo	3.9%
3	YouTube	8.4%	8	Sonico	3.5%
4	LinkedIn	4.6%	9	Fotolog	3.4%
5	SlideShare	4%	10	Google+	2.3%

Tabla 3 – Redes sociales web con mayor porcentaje de aparición

Los porcentajes se obtienen al revisar en cuantos reportes aparece cada red social elegida y normalizando los valores.

A partir de la interacción con las redes sociales del ranking anterior y usando como guía el checklist se obtuvo un listado priorizado de aquellas con mayor grado de colaboración, tal como se muestra en la tabla 4.

#	Red social	% colaboración	#	Red social	% colaboración
1	Google +	100.00%	6	Sonico	63.60%
2	Facebook	90.00%	7	Youtube	63.60%
3	SlideShare	81.80%	8	Twitter	59.10%
4	LinkedIn	77.30%	9	Badoo	45.50%
5	My Space	77.30%	10	Fotolog	31.80%

Tabla 4 – Valoración de redes sociales web acorde a su nivel de colaboración

Los porcentajes generados dan cuenta de una razón entre el número de respuestas positivas encontradas con respecto al número total de preguntas. Por ejemplo, en Google+ todas las herramientas colaborativas están presentes, mientras que en Fotolog aproximadamente tres de cada diez herramientas son incorporadas.

Los encargados del ejercicio de co-creación para el modelo de comercialización atendieron la sugerencia de usar la red social google+ obteniendo resultados muy positivos al igual que comentarios valiosos para la construcción colectiva. Algunos de ellos se enuncian seguidamente:

- *Cliff Zintgraff, CEO at DaVinci Minds: “I believe the most important issue is building a structure that allows the individual IP holders (universities, etc.) to maintain control of university policy issues, while agreeing to share resources for commercialization analysis. I think there are selected policy issues, like inventor royalty agreements, that universities will insist on keeping. Universities’ comfort level that their interests are protected will be key to gaining their enthusiastic support”.*
- *Alex Smeets, CEO at Cambridge Innovation Solutions: “Congratulations on starting this new initiative - by working together I think you’ll achieve a lot more than if every organization in the region tried to do its own technology commercialization in isolation. I believe that the key to successful technology commercialization is to develop the best possible understanding of all the potential applications and markets for the technology. This requires creative, multidisciplinary thinking (to explore the possible applications) as well as a very practical, hands-on approach to obtaining information from the market (typically by talking to parties in the market, not by relying on existing market research reports or databases). If you do this well, then the way forward becomes much clearer and you’ll find it much easier to obtain support and investment. I would therefore suggest that you consider developing a joint market exploration capability”.*

Es de anotar que la red social con mayores características que facilitan la co-creación no es necesariamente la más popular (casi de google+ y facebook).

7. Conclusiones

En este artículo se propuso un protocolo para seleccionar una red social acorde a escenarios específicos de aplicación de la co-creación de valor. La idea surge al detectar un vacío en la revisión de literatura. Actualmente las organizaciones que desean hacer co-creación con sus clientes eligen la herramienta tecnológica a usar de forma arbitraria, con las consecuentes dificultades para replicar ejercicios exitosos y evitar aquellos que no arrojaron resultados positivos.

Esta iniciativa se constituye en una herramienta útil para las organizaciones que requieran directrices sobre la forma de facilitar la interacción de participantes heterogéneos como clientes, competidores, equipos de I+D, personal de mercadeo, entre otros. Existe un sinnúmero de beneficios resultantes al integrar las redes sociales, entre ellos la incorporación de clientes y personal externo e interno de una manera efectiva. Lo anterior a partir de la riqueza de las interacciones que ofrecen.

El uso de redes sociales web a nivel empresarial tiene beneficios tales como la ampliación de contactos, el aceleramiento de procesos de negocio, la mejora en las relaciones con los clientes y la facilidad para reclutar personal de alto nivel (Anria, 2009; Molina, 2001),

además de facilitar la innovación a través de funcionalidades usables y atractivas para cualquier usuario (Ormaetxe, 2008).

En un mundo empresarial donde la presión competitiva está en constante aumento, las empresas están tratando de diferenciarse continuamente. La llegada de las tecnologías Web 2.0, como las redes sociales, permitió a las empresas comunicarse e interactuar con los consumidores y usuarios en línea con el fin de recopilar información y realizar tareas de I + D, marketing y ventas (Roch & Mosconi, 2016).

La iniciativa fue validada con un caso real donde se evidenció que el uso de una red social para una construcción colectiva favorece la sinergia entre participantes heterogéneos. Sin duda alguna, en un mundo cambiante y dependiente de la tecnología, las redes sociales propician procesos colectivos, colaborativos e interactivos. Se trata de herramientas poderosas en otros dominios diferentes al ocio y el entretenimiento.

8. Trabajo futuro

Como trabajo futuro se plantea la identificación y recomendación de las redes sociales más apropiadas para usar de acuerdo a escenarios pre-configurados de co-creación. Por ejemplo, las redes masivas como Facebook y Google+ son útiles para hacer co-creación durante la identificación de oportunidades y generación de ideas en un ciclo de innovación, mientras que LinkedIn (la cual posee herramientas de gestión de proyectos) podría ser usada en ejercicios de co-creación asociados a etapas más avanzadas tales como el diseño o las pruebas.

También vale la pena explorar la posibilidad de proponer una nueva red social especializada para la co-creación de valor. Esta nueva red integraría lo mejor de las redes sociales existentes. También podría pensarse en usar el mismo framework que poseen otras redes sociales (por ejemplo, Facebook), siendo éste mejorado a través de complementos orientados a facilitar las interacciones entre interesados.

Otro trabajo en curso es la exploración de otras redes sociales, y la puesta en marcha del protocolo propuesto en diversos ejercicios reales para verificar su efectividad.

Referencias

- Aiello, L. M., & Ruffo, G. (2012). LotusNet: Tunable privacy for distributed online social network services. *Computer Communications*, 35(1), 75–88. doi: 10.1016/j.comcom.2010.12.006
- Althoff, T., Jindal, P., & Leskovec, J. (2017). Online actions with offline impact: How online social networks influence online and offline user behavior. In *Proceedings of the 10th ACM International Conference on Web Search and Data Mining (WSDM 2017)*. Cambridge, UK.
- Anria, S. (2009). The impact of Social Networking 2.0 on organisations. *Electronic Library*, 27(6), 906–918. doi: 10.1108/02640470911004020
- Caldevilla, D. (2010). Las Redes Sociales: Tipología, uso y consumo de las redes 2.0 en la sociedad digital actual. *Documentación de las Ciencias de la Información*, 33, 45–68.

- Cohen, T., & Clemens, B. (2005). Social networks for creative collaboration. In *Proceedings of the 5th conference on Creativity & cognition (C&C'05)*. London, UK. ACM
- ComScore. (2015). Audience Analytics. Retrieved from <http://www.comscore.com/Products/Audience-Analytics>
- Chesbrough, H. (2006). *Open innovation: a paradigm for understanding industrial innovation*. Oxford, UK: Oxford University Press.
- Dall'Asta, L., Marsili, M., & Pin, P. (2012). Collaboration in social networks. In *Proceedings of the National Academy of Sciences of the United States of America, New York, United States, 109*(12), 4395–4400.
- eBizMBA. (2015). Top 15 Most Popular Social Networking Sites 2016. Retrieved from <http://www.ebizmba.com/articles/social-networking-websites>
- Elogia-Ipsosfacto. (2015). VI Estudio de Redes Sociales de IAB Spain Retrieved from <http://www.informeticplus.com/estudio-anual-redes-sociales-2017-iab-spain-elogia>.
- Enkel, E., Gassmann, O., & Chesbrough, H. (2009). Open R&D and open innovation: exploring the phenomenon. *R&D Management*, 39(4), 311–316.
- Gonzalez, L., González, M., Echeverri, J., & Urrego, G. (2013). Method of interaction in open innovation processes incorporating ubiquitous environments and web 2.0 social networks: A baseline architecture. In *Proceedings the 8th Iberian Conference on Information Systems and Technologies (CISTI 2013)*, IEEE.
- González, M., & González, L. (2015). La co-creación como estrategia para abordar la gobernanza de TI en una organización. *RISTI-Revista Ibérica de Sistemas e Tecnologias de Informação* (15), 01–16.
- González, M., González, L., & Urrego, G. (2013). Survey of interaction in web 2.0 social networks and its application to support open innovation processes. In *Proceedings the 8th Iberian Conference on Information Systems and Technologies (CISTI 2013)*, IEEE.
- Halale, S., Gangadharan, G., & Uden, L. (2015). Enhancing Cocreation Using Social Media. *IT Professional*, 17(2), 40–45.
- Lim, S. L., Damian, D., & Finkelstein, A. (2011). StakeSource2. 0: using social networks of stakeholders to identify and prioritise requirements. In *Proceedings of the 33rd international conference on Software engineering (ICSE'2011)*. Honolulu, Hawaii.
- Maamar, Z., & Badr, Y. (2009). Social networks as a service in modern enterprises. In *Proceedings of the International Conference on the Current Trends in Information Technology (CTIT 2009)*. Dubai, United Arab Emirates. IEEE
- Molina, J. L. (2001). *El análisis de redes sociales*. Barcelona: Editorial Bellaterra.
- Murphy, C., Sheth, S., Kaiser, G., & Wilcox, L. (2008). GenSpace: Exploring social networking metaphors for knowledge sharing and scientific collaborative work. In *Proceedings of the 23rd IEEE/ACM International Conference on Automated Software Engineering (ASE 2008)*. L'Aquila, Italy. IEEE.

- Nafría, I. (2008). *Web 2.0. El usuario, el nuevo rey de Internet*. Barcelona: Ediciones Gestión 2000.
- OECD. (2005). *Oslo Manual: Guidelines for collecting and interpreting innovation data* (3ª ed.). Paris: OECD Publishing.
- Ormaetxe, J. I. (2008). Redes sociales e innovación abierta: apuntes críticos. *Telos: Cuadernos de comunicación e innovación* (76), 114–117.
- Prahalad, C., & Ramaswamy, V. (2004). Co-creation experiences: the next practice in value creation. *Journal of interactive marketing*, 18(3), 5–14.
- Raghavan, P. (2002). Social networks: from the Web to the enterprise. *IEEE Internet Computing*, 6(1), 91–94.
- Reinhardt, W., Moi, M., & Varlemann, T. (2009). Artefact-Actor-Networks as tie between social networks and artefact networks. In *Proceedings of the 5th International Conference on Collaborative Computing: Networking, Applications and Worksharing (CollaborateCom 2009)*. Edinburgh, Great Britain. IEEE.
- Requena, S. (2008). El modelo constructivista con las nuevas tecnologías: aplicado en el proceso de aprendizaje. *RUSC. Universities and Knowledge Society Journal*, 5(2), 26–35.
- Roch, J., & Mosconi, E. (2016). The use of social media tools in the product life cycle phases: A systematic literature review. In *Proceedings of the 49th Hawaii International Conference on System Sciences (HICSS 2016)*. Koloa, USA.
- Tavante, H. C. A., & Coutinho, M. P. (2011). *Web social networks meters and business usage analisys*. In *Proceedings of the International Conference on Computational Aspects of Social Networks (CASoN 2011)*. Salamanca, Spain.
- Thrift, N. (2006). Re-inventing invention: new tendencies in capitalist commodification. *Economy and Society*, 35(2), 279–306.
- Tomsic, A., & Suthers, D. (2005). Effects of a Discussion Tool on Collaborative Learning and Social Network Structure within an Organization. In *Proceedings of the 38th Annual Hawaii International Conference on the System Sciences (HICSS '2005)*. Big Island, HI, USA.
- Wang, P., Li, H., & Suomi, R. (2016). Value Co-Creation in Business via Social Media: a Technology Affordance Approach. In *Proceedings of the Pacific Asia Conference on Information Systems (PACIS 2016)*. Chiayi, Taiwan.
- Yang, S. J. H., & Chen, I. Y. L. (2008). A social network-based system for supporting interactive collaboration in knowledge sharing over peer-to-peer network. *International Journal of Human-Computer Studies*, 66(1), 36–50. doi: 10.1016/j.ijhcs.2007.08.005
- Yanqui, F., & Quispe, M. (2015). Usode las redes sociales en el Perú. *Investigación Andina*, 15(2).

Factores sociales y humanos que afectan el proceso de educación de requerimientos: una revisión sistemática

Ítalo Donoso Barraza¹, Vianca Vega Zepeda¹

italo.donoso@ucn.cl, vvega@ucn.cl

¹Departamento de Ingeniería de Sistemas y Computación, Universidad Católica del Norte, Angamos 0610, 1270709, Antofagasta, Chile.

DOI: 10.17013/risti.n.69–83

Resumen: La Ingeniería de Requerimientos (RE, por sus siglas en inglés) es uno de los procesos más importantes durante el desarrollo de software y su correcta ejecución afecta en gran medida el éxito del producto desarrollado. El objetivo general de este proceso es entender con precisión el problema que se pretende resolver. Las actividades de la RE requieren un alto grado de interacción y comunicación entre todos los involucrados. Múltiples factores afectan este proceso, destacando el aspecto social y emocional de la persona, que no son considerados comúnmente por técnicas de educación tradicionales. El objetivo de este trabajo es identificar factores sociales y humanos no cubiertos durante el proceso de educación y que puedan afectar negativamente el éxito del producto. Lo anterior, mediante un proceso de revisión sistemática sobre fuentes bibliográficas reconocidas por la comunidad científica siguiendo el enfoque de tres grandes actividades descritas por Barbara Kitchenham.

Palabras-clave: Ingeniería de Requerimientos, Factores Sociales, Educación.

Social and human factors that affect the requirements elicitation process: a systematic review

Abstract: Requirements engineering (RE) is one of the most important phases during the software development and its proper execution greatly affects the developed product success. The overall objective of the RE is to understand accurately the problem that it is intended to solve. RE activities require a high degree of interaction and communication between all stakeholders. Multiple factors affect this process, highlighting social and emotional aspects of people, which are not commonly considered by traditional elicitation techniques. The objective of this work is to identify social and human factors that are not covered during the elicitation process and they may negatively affect the product success. The review process was carried out on bibliographic sources recognized by scientific community following three major activities described by Barbara Kitchenham approach.

Keywords: Engineering Requirements, Social Factors, Elicitation.

1. Introducción

La Ingeniería de Requerimientos (RE, por sus siglas en inglés) es un proceso que abarca un conjunto de actividades que se realizan comúnmente al inicio del ciclo de desarrollo de software. La educación de los requerimientos, que tiene por objetivo descubrir las necesidades de los usuarios y capturar los requerimientos del sistema, es la etapa más importante de este proceso y su correcta aplicación repercute en gran medida en el éxito del producto final (Hastie & Wojewoda, 2015). Por otra parte, una incorrecta aplicación de las tareas de educación afectará negativamente en aspectos tan importantes como: el alcance del proyecto, el presupuesto definido y los plazos de entrega acordados previamente (Kumari & Anitha, 2016; Okewu, 2015).

Los desafíos que enfrenta la RE son variados y una gran parte de estos se presentan al momento de educir las necesidades del producto debido a la falta de comunicación e interacción entre el equipo de desarrollo y los usuarios; individuos de distintas disciplinas con diferentes niveles de conocimiento técnico (Hadar, Soffer, & Kenzi, 2014). La Ingeniería de Requerimientos debe lidiar con la brecha de comunicación entre todos los involucrados. No solo debe enfocarse en el medio utilizado para producir la comunicación, o los artefactos dispuestos en las instancias en que se reúnen stakeholders y especialistas, es igual de importante considerar aspectos cognitivos, emocionales y motivacionales propios de la «naturaleza social» de la persona que contribuyan a la transferencia del conocimiento de manera exitosa (Apshvalka, Donina, & Kirikova, 2009). El conocimiento es la fuente para definir y estructurar los requerimientos de un producto software, este puede provenir del ambiente o dominio de la aplicación, estar alineado con las estrategias del negocio, encontrarse inmerso en las infraestructuras organizacionales, o suscitado en actividades financieras y operacionales del negocio (Burnay, 2016). Sin embargo, la fuente más importante es el individuo y su conocimiento reside en la mente humana.

Si bien existen diversas técnicas para elicitar los requerimientos de software en un contexto particular (Carrizo, Dieste, & Juristo, 2014), no existe un conjunto amplio de estudios que abarquen aspectos sociales y humanos en el proceso de educación. En este sentido, el objetivo del presente trabajo es identificar qué aspectos sociales propios de la naturaleza de la persona no son considerados comúnmente cuando se realizan tareas de educación de requerimientos que pueden afectar negativamente el éxito del producto final. La identificación de dichos aspectos se realizó mediante un proceso de revisión de la literatura siguiendo el enfoque de Bárbara Kitchenham (Kitchenham, 2004) sobre cuatro fuentes bibliográficas reconocidas por la comunidad científica. Los artículos seleccionados corresponden a trabajos ligados al ámbito de la Ingeniería de Software y, en específico, a la Ingeniería de Requerimientos enfocados en el proceso de educación de requerimientos sobre un producto software. El trabajo realizado y los resultados descritos en dichos artículos fueron analizados, agrupados y clasificados en seis categorías llamados “factores sociales y humanos”.

El presente documento se estructura como sigue: la sección 2 describe de manera general el método utilizado para ejecutar la revisión sistemática, la sección 3 describe el proceso de planificación previo a la búsqueda de los estudios, la sección 4 muestra el método empleado para la selección de los artículos a considerar dentro de la revisión, la sección

5 sintetiza los resultados extraídos de los veinticinco artículos analizados, la sección 6 enfatiza las conclusiones del trabajo realizado orientado a los resultados obtenidos, y finalmente la sección 7 describe la literatura consultada.

2. Proceso de investigación

Los objetivos que guían la ejecución de una revisión sistemática son variados y están conectados estrechamente con el tema de investigación abordado (Kitchenham, 2004). Sin embargo, el foco principal de una revisión sistemática es sintetizar el trabajo existente dentro de un área de interés en particular, mediante una estrategia de búsqueda definida que permita evaluar la completitud de la estudios analizados (Kitchenham, 2004).

La revisión sistemática se realizó siguiendo tres grandes actividades: planificación, ejecución y reporte de los resultados obtenidos según el enfoque de Barbara Kitchenham (Kitchenham, 2004). Cada una de estas actividades a su vez describen un conjunto de tareas relevantes que se realizan dependiendo del foco u objetivo del estudio. Para la presente revisión sistemática se han adaptado ciertas tareas descritas por dichas actividades. Las siguientes secciones describen las actividades realizadas en cada una de las actividades mencionadas anteriormente, es decir: planificación, ejecución y reporte.

3. Planificación

3.1. Pregunta de investigación

En cualquier revisión sistemática existe una o más preguntas de investigación que motivan el estudio de un tema particular en un área de investigación específica. Con la ejecución de la revisión sistemática se pretende dar respuesta a dicha pregunta mediante el análisis y síntesis de los datos obtenidos. De este modo, el área de investigación corresponde a la Ingeniería de Software y más en particular, al proceso de educación de requerimientos dentro de la RE.

La pregunta de investigación que centró la realización del presente artículo de revisión fue:

RQ1: ¿Qué problemas y/o factores sociales y humanos afectan el proceso de obtención de requerimientos de software?

3.2. Lenguaje del estudio

Dentro del universo de estudios analizados solo se consideraron aquellos artículos escritos en el idioma inglés o español.

3.3. Palabras claves y sinónimos

En la Tabla 1 se especifican las palabras claves que fueron consideradas en la revisión sistemática y que, a su vez, fueron utilizadas para formular la cadena de búsqueda.

Palabra en inglés	Palabra en español
Problem	Problema
Issue	Problema
Factors	Factores
Human	Humano
Social	Social
Elicitation	Elicitación
Requirements	Requerimientos
Engineering	Ingeniería
Software	Software

Tabla 1 – Palabras clave

Se ha utilizado la palabra «elicitación» en lugar de «educación» en la cadena de búsqueda en español. Lo anterior, debido a que, dentro de la Ingeniería de Software y dentro del contexto del idioma inglés, se utiliza para describir la transferencia de información de un individuo a otro mediante artefactos, técnicas o metodologías definidas con el objetivo de capturar requerimientos y descubrir necesidades de un cliente. Se entiende por individuo a uno o más grupos de personas, o entidades que conforman una organización.

3.4.Cadena de búsqueda

La cadena utilizada en los motores de búsqueda de las revistas científicas se puede observar en la Tabla 2. Al igual que las palabras clave, se ha utilizado la cadena de búsqueda tanto en inglés como en español. Es importante destacar que para algunos motores no está admitido el uso de caracteres especial como paréntesis, debiendo adaptar así la cadena a la sintaxis del buscador en cuestión.

Cadena en inglés	Cadena en español
("PROBLEMS" OR "ISSUES") AND (("SOCIAL" OR "HUMAN") AND "FACTORS") AND (("REQUIREMENTS ENGINEERING") OR "ELICITATION") AND "SOFTWARE"	("PROBLEMAS" Y ("FACTORES" Y (("SOCIALES" O "HUMANOS") Y ("INGENIERÍA DE REQUERIMIENTOS" O "ELICITACION") Y "SOFTWARE"

Tabla 2 – Cadenas de búsqueda

3.5.Fuentes bibliográficas

Los artículos se han obtenido a partir de cuatro fuentes de investigación reconocidas por la comunidad científica. El acceso a los recursos ha sido posible mediante los privilegios otorgados a la cuenta de alumno de la Universidad Católica del Norte y su vínculo con la fuente.

A continuación, se indica el listado de fuentes consultadas:

- ACM Digital Library
- Science@Direct
- SpringerLink
- ISI Web of Knowledge

4. Ejecución

Una vez definida la cadena de búsqueda y las fuentes bibliográficas a consultar, se procedió a obtener los artículos desde las revistas científicas mediante la aplicación de la cadena en los motores disponibles en el sitio web de cada fuente.

Los trabajos recuperados fueron ordenados siguiendo los siguientes filtros:

- Año de publicación comprendido entre 2011 y 2017
- Orden descendente según la relevancia del artículo (de acuerdo al número de coincidencias de las palabras clave contenidas en la cadena de búsqueda y aquellas mencionadas en el cuerpo del artículo)
- Área de interés relacionada con «Ciencias de la Computación» e «Ingeniería de Software»
- Idioma de los artículos en inglés o español

La cantidad de artículos recuperados, luego de aplicar los filtros, se muestran en la Tabla 3. Como se puede observar, la gran mayoría de artículos recuperados comprenden las fuentes bibliográficas Science@Direct y SpringerLink alcanzando un 98% del total de estudios obtenidos entre ambas fuentes. La totalidad de los resultados arrojados por los motores de las fuentes comprendieron trabajos publicados en el idioma inglés, no así en español para el cual no se obtuvieron artículos.

Fuente	Cantidad de documentos en inglés	Distribución porcentual respecto al total de artículos
ACM Digital Library	346	2.4%
Science@Direct	8540	58.5%
SpringerLink	5691	39%
ISI Web of Knowledge	11	0.1%
Totales	14588	100%

Tabla 3 – Cantidad de artículos obtenidos

4.1. Criterios de inclusión y exclusión

Los criterios de inclusión y exclusión deben estar alineados con el objetivo de la revisión sistemática. Estos criterios son aplicados con el fin de filtrar y seleccionar aquellos estudios primarios que proporcionen una evidencia directa y que permitan responder a la pregunta de investigación definida (Kitchenham, 2004).

La Tabla 4 resume los criterios de inclusión y exclusión considerados en la ejecución de esta revisión sistemática.

N.º	Criterio de inclusión	Criterio de exclusión
1	La fecha del artículo es superior o igual a 2011	La fecha del artículo es inferior a 2011
2	El artículo se encuentra en inglés o español	El artículo no se encuentra en inglés o español
3	Es un artículo de revista	No es un artículo de revista
4	El artículo contiene en su título palabras claves definidas en la cadena de búsqueda	El artículo no tiene contiene en su título palabras claves definidas en la cadena de búsqueda
5	El artículo menciona en su resumen palabras claves definidas en la cadena de búsqueda	El artículo no menciona en su resumen palabras claves definidas en la cadena de búsqueda
6	El artículo incluye descripciones o referencias en el cuerpo del mismo asociadas a las palabras claves definidas en la cadena de búsqueda	El artículo no incluye descripciones o referencias en el cuerpo del mismo asociadas a las palabras claves definidas en la cadena de búsqueda
7	No ha sido seleccionado anteriormente	Ya ha sido seleccionado anteriormente
8	Es un artículo de acceso gratuito	No es un artículo de acceso gratuito
9	Es un artículo accesible a través de cuenta de alumno de la Universidad Católica del Norte	No es un artículo accesible a través de cuenta de alumno de la Universidad Católica del Norte

Tabla 4 – Criterios de inclusión y exclusión

4.2.Procedimiento para la selección de artículos

El primer paso consistió en adaptar la cadena de búsqueda a la sintaxis definida de los cuatro motores virtuales de las fuentes bibliográficas. Luego, se aplicaron los filtros por año y área (ciencias de la computación e ingeniería de software) y se ordenaron de acuerdo al nivel de ocurrencia de las palabras claves tanto en el título, como en el resumen (abstract).

Posteriormente, fueron seleccionados los primeros 40 artículos de cada revista a los cuales se les aplicó los criterios de inclusión y exclusión definidos en la Tabla 4. En la fuente ISI Web of Knowledge el buscador arrojó solo 11 artículos los cuales fueron considerados en su totalidad para aplicar los criterios antes mencionados. En ningún motor de búsqueda se encontraron documentos en español.

Una vez aplicados los criterios de inclusión, se logró obtener un total de 25 artículos elegibles para someterlos al proceso de análisis de datos. La Tabla 5 muestra la cantidad de artículos seleccionados por fuente bibliográfica.

Fuente	Cantidad de documentos en inglés	Distribución porcentual respecto al total de artículos
ACM Digital Library	2	8%
Science@Direct	5	20%
SpringerLink	17	68%
ISI Web of Knowledge	1	4%
Totales	25	100%

Tabla 5 – Cantidad de artículos seleccionados por fuente

Los estudios se clasificaron con base en su objetivo principal agrupándolos en dos categorías generales:

1. Estudios teóricos, empíricos y propuestas metodológicas enfocadas en el proceso de educación y,
2. Aquellos que incorporan y sustentan la aplicación del punto 1 en uno o más casos de estudios.

La Tabla 6 dispone los veinticinco artículos seleccionados, entre los cuales un 36% de ellos se basan en la aplicación mediante casos de estudios referentes al proceso de Ingeniería de Requerimientos. Por un tema de legibilidad, se han suprimido los nombres de los artículos. Sin embargo, estos se encuentran debidamente referenciados en el apartado final de este documento.

Los artículos fueron analizados enfocándose en el proceso de educación considerando el factor humano como eje central, destacando los inconvenientes o problemas al momento de educir, validar y documentar los requerimientos según el desarrollo involucrado. Como resultado, se agruparon en seis categorías los atributos recurrentes y relacionados al ámbito social y humano del proceso: ambigüedad, conocimiento tácito, comunicación, entendimiento, ámbito cognitivo, emocional y motivacional, resistencia y valores humanos.

5. Resultados y discusión

La Ingeniería de Requerimientos es un proceso complejo y a la vez crítico durante el desarrollo de software. La información capturada por los Ingenieros o Especialistas en Requerimientos debe ser interpretada, analizada, modelada y fuertemente validada (Hadar et al., 2014) antes de comenzar las etapas propias del ciclo de desarrollo, independiente del modelo tradicional o ágil que se utilice (Würfel et al., 2016).

La etapa de educación de los requerimientos es una de las actividades de comunicación más intensas de este proceso la cual demanda una compleja interacción entre los individuos que se extiende desde el entendimiento de las necesidades del negocio, la toma de decisiones, la gestión de los cambios y la resolución de conflictos entre los stakeholders (Calefato et al., 2012). Si bien la comunicación es el núcleo para producir especificaciones sólidas que se sustenten en los datos capturados, se requiere adicionalmente un canal o medio que permita que esta fluya a través de los distintos stakeholders dentro de la organización que estén interesados en el producto a desarrollar (Liebel et al., 2016).

Cód.	Autores	Tipo
P01	(Ferrari, Spoletini, & Gnesi, 2016)	Caso de estudio
P02	(Liebel, Tichy, Knauss, Ljungkrantz, & Stieglbauer, 2016)	Caso de estudio
P03	(Calefato, Damian, & Lanubile, 2012)	Empírico
P04	(Apshvalka, Donina, & Kirikova, 2009)	Empírico
P05	(Sethia & Pillai, 2014)	Empírico
P06	(Burnay, Jureta, & Faulkner, 2014a)	Empírico Teórico
P07	(Harbers, Detweiler, & Neerinx, 2015)	Empírico Caso de estudio
P08	(Okewu, 2015)	Caso de estudio
P09	(Kumari & Anitha, 2016)	Empírico Teórico
P10	(Ramachandran, Dodda, & Santapoor, 2011)	Empírico
P11	(Anwar, Razali, & Ahmad, 2011)	Empírico Teórico
P12	(Pa & Zin, 2011)	Caso de estudio
P13	(Gervasi et al., 2013)	Empírico Teórico
P14	(Shuhud, Richter, & Ahmad, 2013)	Caso de estudio
P15	(Fuentes-Fernández, Gómez-Sanz, & Pavón, 2010)	Caso de estudio
P16	(Katina, Keating, & Jaradat, 2014)	Empírico
P17	(Hadar, I., Soffer, P., & Kenzi, 2014)	Empírico
P18	(Bjarnason et al., 2014)	Caso de estudio
P19	(McLeod & MacDonell, 2011)	Empírico
P20	(Burnay, 2016)	Empírico Teórico
P21	(Dey & Lee, 2017)	Empírico Teórico
P22	(Burnay, Jureta, & Faulkner, 2014b)	Empírico Teórico
P23	(Azadegan, Papamichail, & Sampaio, 2013)	Caso de estudio
P24	(Würfel, Lutz, & Diehl, 2016)	Empírico Teórico
P25	(Lopes & Forster, 2013)	Empírico Teórico

Tabla 6 – Artículos seleccionados

La entrevista es el método empleado mayoritariamente en el área y el entrevistador debe poseer la capacidad de hacer frente a las «emociones» de los stakeholders para poder manejarlos, promover la motivación en la participación, procurar que la comunicación se realice «cara a cara» transmitiendo confianza, seguridad y describiendo ventajas

para el usuario sobre el nuevo producto desarrollado o adquirido (Apshvalka et al., 2009). Lo anterior, es una tarea compleja ya que, en este escenario, el especialista necesita desempeñarse como motivador y «psicólogo» a la vez para poder capturar las necesidades de los individuos de la manera correcta. Esto conlleva la comprensión de procesos cognitivos, emocionales y motivacionales de quienes están inmersos en el proceso desde la perspectiva del cliente y del equipo de desarrollo; tarea compleja especialmente para un ingeniero de software cuya formación profesional excluye este ámbito. Estas múltiples técnicas de educación no contemplan ampliamente estos aspectos sociales a priori considerando que pueden afectar negativamente el éxito del proyecto si no se controlan de la manera adecuada.

Como se ha mencionado anteriormente, la RE es un proceso intensivo de transferencia de información entre quienes son «poseedores del conocimiento» desde el lado del cliente-usuario y el equipo de desarrollo. Este conocimiento está incrustado casi en su mayoría en los stakeholders y es labor del especialista en requerimientos poder abstraer este conocimiento apoyado por procesos cognitivos, emocionales y motivacionales (Apshvalka et al., 2009). En este sentido, producto de la ejecución de la presente revisión sistemática, se identificaron aquellos factores sociales y humanos provenientes desde analistas y clientes que pueden afectar el proceso de educación y, en contra parte, otorgar una base para su consideración al momento de efectuar actividades de RE. La ocurrencia de los problemas y/o inconvenientes que exponen los artículos al momento de educir requerimientos fueron agrupados en una categoría de más alto nivel que responde a factores sociales y humanos descritos en la Tabla 7.

Factores sociales y humanos (categoría)	P01	P02	P03	P04	P05	P06	P07	P08	P09	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	P25
Ambigüedad	x										x	x			x	x	x			x		x		x	
Conocimiento tácito				x		x							x				x					x			
Comunicación		x	x	x	x				x	x	x	x	x	x	x		x	x	x	x		x		x	x
Entendimiento					x				x	x	x	x		x					x						
Ámbito cognitivo, emocional y motivacional				x		x						x				x									
Resistencia				x							x					x									
Valores humanos							x				x														

Tabla 7 – Factores sociales y humanos detectados

A continuación, se destacan los aspectos más relevantes que definen a cada factor descrito en la Tabla 7.

Ambigüedad

Se traduce comúnmente en malentendidos durante las entrevistas provocando especificaciones poco claras e incompletas producto del uso del lenguaje natural. Desde el lado del cliente se manifiesta en el sentido de que este último se expresa utilizando un lenguaje lleno de jergas asociadas a su dominio provocando la interpretación múltiple de los analistas (Burnay et al., 2014a), o entrega muy poca o demasiada información que se aleja de los límites de procesamiento del analista durante una entrevista.

Conocimiento tácito

El conocimiento tácito dentro de la RE se refiere al «conocimiento» que no es transmitido desde el cliente al analista (Ferrari et al., 2016). Este conocimiento tácito es difícil de adquirir si el analista no se encuentra inmerso dentro del dominio de los usuarios y/o del cliente. En general, la psicología indica que tener conocimiento previo del dominio puede causar tendencia a trabajar de la misma forma y abordar los problemas bajo esa perspectiva sin proceder a un análisis previo; es un inhibidor de la creatividad (Hadar et al., 2014).

Comunicación

La comunicación es sin duda el elemento intrínseco del proceso de educación, sin un adecuado entendimiento de las partes existe un alto índice de rechazo del producto final dados los vacíos y malentendidos durante el proceso de análisis. Los principales factores que afectan la comunicación se producen desde el lado del cliente y abarcan distintos ámbitos:

- Comunicación infrecuente producto de las barreras organizacionales (Sethia & Pillai, 2014).
- Falta de habilidades de presentación, habilidades de escritura, falta en la organización de ideas y organización general de la información.
- Respuesta tardía con base en el canal de comunicación utilizado aumentando los tiempos que tomará el desarrollo del producto.

Las habilidades de comunicación de los stakeholders determinarán la calidad de los requerimientos (Shuhud et al., 2013).

Entendimiento

El entendimiento mutuo está estrechamente ligado con los tres factores descritos anteriormente, es decir, ambigüedad, conocimiento tácito y comunicación que producen generalmente malentendidos. Esto se debe a que la RE es un proceso en donde participan individuos de distintas disciplinas abarcando un amplio conjunto de conceptos propios a cada una de ellas. El foco central del entendimiento es la negociación la cual es fundamental para definir los límites del sistema, establecer prioridades en los requerimientos, identificar aquellas personas relevantes con la autoridad y poder

en la toma de decisiones, y finalmente discutir los cambios que se producen sobre los requerimientos (Anwar et al., 2011; Shuhud et al., 2013). El especialista en requerimientos debe poseer un alto grado de comunicación interpersonal que le permita negociar en situaciones críticas y, a su vez, tratar con las diversas personalidades de los stakeholders enfocado en la resolución de conflictos.

Ámbito cognitivo, emocional y motivacional

La dimensión social es un área muy poco explorada dentro de la Ingeniería de Software, esto se debe a que los ingenieros se enfocan más bien en aspectos técnicos por sobre los sociales (Ramachandran et al., 2011). En esta área destacan tres aspectos importantes (Apshvalka et al., 2009):

1. **Cognitivo:** los individuos hacen uso de sus facultades mentales para producir la interacción analistas y stakeholders, esto les permite percibir la información de su entorno, interpretarla, representarla para el uso correcto en la toma de decisiones.
2. **Emocional:** permite reconocer las «emociones» de los stakeholders derivadas de lo que supone representará el sistema para su entorno o trabajo. En este ámbito es importante que el analista disponga de las cualidades que le permitan responder ante situaciones en donde los stakeholders, por ejemplo, asuman que el nuevo producto supone una carga adicional en su forma de trabajar, o en un caso extremo, que dicho producto reemplazará las actividades que desempeña mediante procesos automatizados de software. En ambos casos el especialista en requerimientos debe elegir cuidadosamente el método y forma en que expone sus inquietudes con todos los involucrados.
3. **Motivacional:** es fundamental para predecir, detectar y evitar emociones reacias durante el proceso de educación; el ingeniero debe actuar como un pseudo-sicólogo promoviendo la participación del proceso en un ambiente agradable, de confianza, franqueza y empatía.

Resistencia

La resistencia al cambio es un reto conocido dentro la Ingeniería de Software y en particular en la RE se traduce en la falta de participación de los stakeholders en las actividades de educación (Anwar et al., 2011). Este desinterés está directamente conectado con el ámbito emocional y motivacional de la persona, ya que si el analista no expone de manera directa los beneficios del nuevo producto para los stakeholders, estos últimos no se verán atraídos por el proceso, provocando finalmente que el analista seleccione o consulte a otros individuos con un rango menor en la jerarquía organizacional lo cual conllevará a un inminente cambio en la especificación que puede suscitarse en fases de implementación y despliegue (Apshvalka et al., 2009).

Valores humanos

Un aspecto que no se considera prácticamente en los proyectos de desarrollo es la valoración que los usuarios poseen del sistema. Esta valoración abarca la confianza, autonomía y seguridad que el sistema le brinda al usuario ya sea en su trabajo o en

su vida cotidiana (Harbers et al., 2015). Estos no siempre son previsibles durante el desarrollo del producto ya que un producto muy seguro restaría facilidad de uso y, en el otro extremo, un producto eficiente no necesariamente sería confiable. Actualmente es un campo que posee muy poca evidencia teórica y empírica que abarque la importancia que supone la existencia de un producto software dentro del entorno de trabajo de los usuarios alineado con lo que estos últimos creen importante y valioso dentro de su entorno, vida e interacción con sus pares.

6. Conclusiones

El proceso de educación de requerimientos varía entre una técnica y otra, y en algunos casos está alineado con el enfoque tradicional o ágil que se emplee para educir las necesidades de los stakeholders. Independientemente de la técnica utilizada, los retos que enfrenta la RE son variados y existe un porcentaje de estos asociados a la naturaleza social de la persona que no son tratados a priori por técnicas de educación existentes. Lo anterior, se debe a que el Ingeniero de Software no recibe una formación profesional enfocada en el aspecto social y humano dentro de su disciplina, lo que provoca que sea un área difícil de abordar, más no imposible, desde la perspectiva de las ciencias de la ingeniería.

La comunicación es, sin lugar a dudas, el núcleo de las actividades definidas por la Ingeniería de Requerimientos considerando que en el proceso participan individuos con preferencias, prioridades, objetivos y metas distintas del dominio que hacen que los especialistas en requerimientos necesiten de enfoques más amplios que les entreguen las herramientas necesarias para cubrir, en parte, las brechas sociales que se producen en las etapas tempranas de desarrollo.

La identificación de estos factores entrega una base conceptual respecto a lo que se debería considerar al momento de establecer actividades de comunicación con los stakeholders con el objetivo de utilizar, por ejemplo, herramientas que promuevan una comunicación fluida eliminando espacios de ambigüedad y resistencia, favoreciendo el entendimiento entre ambas partes y la transmisión del conocimiento que no se entrega explícitamente (conocimiento tácito). Estos factores sirven como base inicial para investigar otros temas respecto a cómo capturar de manera simple y óptima el conocimiento tácito durante las entrevistas u otras instancias de reunión, la preparación o entrenamiento en el aspecto emocional y motivacional que debería desarrollar el especialista en requerimientos y que le permita tratar con los distintos caracteres que podrían manifestar los stakeholders. De este último párrafo se desprende que un especialista en requisitos debería poseer las aptitudes, adicionales a su disciplina, que sean necesarias para realizar actividades de educación relacionadas al factor humano.

Estas aptitudes deberían incluir:

- Habilidades en la adquisición del conocimiento tácito relevante dentro del dominio del producto ya sea de manera cognitiva o apoyada por un artefacto concreto.
- Habilidades propias de la interacción social mediante la comunicación verbal y escrita.

- Grandes capacidades de negociación, priorización y toma de decisiones respecto a los stakeholders.
- La generación de especificaciones simples, estructuradas, objetivas y sin ambigüedades.
- Uso y manejo de herramientas que favorezcan la participación y motivación de los distintos involucrados en el proceso disminuyendo la resistencia de estos.
- Describir y extender en todo momento la creación del software como un producto que entregará seguridad y confianza a los stakeholders en las tareas que realizan dentro de su área.

Referencias

- Anwar, F., Razali, R., & Ahmad, K. (2011). Achieving effective communication during requirements elicitation - A conceptual framework. *Communications in Computer and Information Science*, 181, 600–610. https://doi.org/10.1007/978-3-642-22203-0_51
- Apshvalka, D., Donina, D., & Kirikova, M. (2009). Understanding the Problems of Requirements Elicitation Process: A Human Perspective. In W. Wojtkowski, G. Wojtkowski, M. Lang, K. Conboy, & C. Barry (Eds.), *Information Systems Development* (pp. 211–223). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-68772-8_17
- Azadegan, A., Papamichail, K. N., & Sampaio, P. (2013). Applying collaborative process design to user requirements elicitation: A case study. *Computers in Industry*, 64(7), 798–812. <https://doi.org/https://doi.org/10.1016/j.compind.2013.05.001>
- Bjarnason, E., Runeson, P., Borg, M., Unterkalmsteiner, M., Engström, E., Regnell, B., Feldt, R. (2014). Challenges and practices in aligning requirements with verification and validation: a case study of six companies. *Empirical Software Engineering*, 19(6), 1809–1855. <https://doi.org/10.1007/s10664-013-9263-y>
- Burnay, C. (2016). Are Stakeholders the Only Source of Information for Requirements Engineers? Toward a Taxonomy of Elicitation Information Sources. *ACM Trans. Manage. Inf. Syst.*, 7(3), 8:1–8:29. <https://doi.org/10.1145/2965085>
- Burnay, C., Jureta, I. J., & Faulkner, S. (2014a). An Exploratory Study of Topic Importance in Requirements Elicitation Interviews. In M. Jarke, J. Mylopoulos, C. Quix, C. Rolland, Y. Manolopoulos, H. Mouratidis, & J. Horkoff (Eds.), *Lecture Notes in Computer Science: Vol. 8484. Advanced Information Systems Engineering* (pp. 180–195). https://doi.org/10.1007/978-3-319-07881-6_13
- Burnay, C., Jureta, I. J., & Faulkner, S. (2014b). What stakeholders will or will not say: A theoretical and empirical study of topic importance in Requirements Engineering elicitation interviews. *Information Systems*, 46, 61–81. <https://doi.org/https://doi.org/10.1016/j.is.2014.05.006>
- Calefato, F., Damian, D., & Lanubile, F. (2012). Computer-mediated communication to support distributed requirements elicitations and negotiations tasks. *Empirical Software Engineering*, 17(6), 640–674. <https://doi.org/10.1007/s10664-011-9179-3>

- Carrizo, D., Dieste, O., & Juristo, N. (2014). Systematizing requirements elicitation technique selection. *Information and Software Technology*, 56(6), 644–669. <https://doi.org/10.1016/j.infsof.2014.01.009>
- Dey, S., & Lee, S.-W. (2017). REASSURE: Requirements elicitation for adaptive socio-technical systems using repertory grid. *Information and Software Technology*, 87, 160–179. <https://doi.org/https://doi.org/10.1016/j.infsof.2017.03.004>
- Ferrari, A., Spoletini, P., & Gnesi, S. (2016). Ambiguity and tacit knowledge in requirements elicitation interviews. *Requirements Engineering*, 21(3), 333–355. <https://doi.org/10.1007/s00766-016-0249-3>
- Fuentes-Fernández, R., Gómez-Sanz, J. J., & Pavón, J. (2010). Understanding the human context in requirements elicitation. *Requirements Engineering*, 15(3), 267–283. <https://doi.org/10.1007/s00766-009-0087-7>
- Gervasi, V., Gacitua, R., Rouncefield, M., Sawyer, P., Kof, L., Ma, L., ... Nuseibeh, B. (2013). Unpacking Tacit Knowledge for Requirements Engineering. In W. Maalej & A. Thurimella (Eds.), *Managing Requirements Knowledge* (pp. 23–47). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-34419-0_2
- Hadar, I., Soffer, P., & Kenzi, K. (2014). The role of domain knowledge in requirements elicitation via interviews: an exploratory study. *Requirements Engineering*, 19(2), 143–159. <https://doi.org/10.1007/s00766-012-0163-2>
- Harbers, M., Detweiler, C., & Neerincx, M. A. (2015). Embedding Stakeholder Values in the Requirements Engineering Process. In S. Fricker & K. Schneider (Eds.), *Lecture Notes in Computer Science: Vol. 9013. Requirements Engineering: Foundation for Software Quality* (pp. 318–332). https://doi.org/10.1007/978-3-319-16101-3_23
- Hastie, S., & Wojewoda, S. (2015). *Standish Group 2015 Chaos Report - Q&A with Jennifer Lynch*. Retrieved from <https://www.infoq.com/articles/standish-chaos-2015>
- Katina, P. F., Keating, C. B., & Jaradat, R. M. (2014). System requirements engineering in complex situations. *Requirements Engineering*, 19(1), 45–62. <https://doi.org/10.1007/s00766-012-0157-0>
- Kitchenham, B. (2004). *Procedures for performing systematic reviews*. (Report TR/SE-0401). Keele, UK: Keele University.
- Kumari, S. N., & Anitha, S. P. (2016). An Extended Study on the Association Between Elicitation Issues and Software Project Performance: A Theoretical Model. In V. Snášel, A. Abraham, P. Krömer, M. Pant, & A. Muda (Eds.), *Innovations in Bio-Inspired Computing and Applications: Vol. 424. Advances in Intelligent Systems and Computing* (pp. 523–533). https://doi.org/10.1007/978-3-319-28031-8_46
- Liebel, G., Tichy, M., Knauss, E., Ljungkrantz, O., & Stieglbauer, G. (2016). Organisation and communication problems in automotive requirements engineering. *Requirements Engineering*. <https://doi.org/10.1007/s00766-016-0261-7>

- Lopes, M. E. R. F., & Forster, C. H. Q. (2013). Application of human error theories for the process improvement of Requirements Engineering. *Information Sciences*, 250, 142–161. <https://doi.org/https://doi.org/10.1016/j.ins.2013.07.010>
- McLeod, L., & MacDonell, S. G. (2011). Factors that affect software systems development project outcomes. *ACM Computing Surveys*, 43(4), 1–56. <https://doi.org/10.1145/1978802.1978803>
- Okewu, E. (2015). Requirements Engineering in an Emerging Market. In O. Gervasi, B. Murgante, S. Misra, M. L. Gavrilova, A. M. A. C. Rocha, C. Torre, ... B. O. Apduhan (Eds.), *Lecture Notes in Computer Science: Vol. 9158 Computational Science and Its Applications – ICCSA 2015* (pp. 476–491). https://doi.org/10.1007/978-3-319-21410-8_37
- Pa, N. C., & Zin, A. M. (2011). Managing Communications Challenges in Requirement Elicitation. In J. M. Zain, W. M. bt Wan Mohd, & E. El-Qawasmeh (Eds.), *Software Engineering and Computer Systems* (pp. 803–811). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-22203-0_67
- Ramachandran, S., Dodda, S., & Santapoor, L. (2011). Overcoming Social Issues in Requirements Engineering. In N. Meghanathan, B. K. Kaushik, & D. Nagamalai (Eds.), *Advanced Computing* (pp. 310–324). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-17881-8_30
- Sethia, N. K., & Pillai, A. S. (2014). The Effects of Requirements Elicitation Issues on Software Project Performance: An Empirical Analysis. In C. Salinesi & I. van de Weerd (Eds.), *Lecture Notes in Computer Science: Vol. 8396. Requirements Engineering: Foundation for Software Quality. REFSQ 2014* (pp. 285–300). https://doi.org/10.1007/978-3-319-05843-6_21
- Shuhud, M. I. M., Richter, A., & Ahmad, A. (2013). Supporting Requirements Elicitation Practices. In P. Antunes, M. A. Gerosa, A. Sylvester, J. Vassileva, & G.-J. de Vreede (Eds.), *Lecture Notes in Computer Science Vol. 8224. Collaboration and Technology. CRIWG 2013* (pp. 306–321). https://doi.org/10.1007/978-3-642-41347-6_22
- Wüffel, D., Lutz, R., & Diehl, S. (2016). Grounded requirements engineering: An approach to use case driven requirements engineering. *Journal of Systems and Software*, 117, 645–657. <https://doi.org/10.1016/j.jss.2015.10.024>

Crítérios Editoriais

A RISTI (Revista Ibérica de Sistemas e Tecnologias de Informação) é um periódico científico, propriedade da AISTI (Associação Ibérica de Sistemas e Tecnologias de Informação), que foca a investigação e a aplicação prática inovadora no domínio dos sistemas e tecnologias de informação.

O Conselho Editorial da RISTI incentiva potenciais autores a submeterem artigos originais e inovadores para avaliação pelo Conselho Científico.

A submissão de artigos para publicação na RISTI deve realizar-se de acordo com as chamadas de artigos e as instruções e normas disponibilizadas no sítio Web da revista (<http://www.risti.xyz/>).

Todos os artigos submetidos são avaliados por um conjunto de membros do Conselho Científico, não inferior a três elementos.

Em cada número da revista são publicados entre cinco a oito dos melhores artigos submetidos.

Criterios Editoriales

La RISTI (Revista Ibérica de Sistemas y Tecnologías de la Información) es un periódico científico, propiedad de la AISTI (Asociación Ibérica de Sistemas y Tecnologías de la Información), centrado en la investigación y en la aplicación práctica innovadora en el dominio de los sistemas y tecnologías de la información.

El Consejo Editorial de la RISTI incentiva autores potenciales a enviar sus artículos originales e innovadores para evaluación por el Consejo Científico.

El envío de artículos para publicación en la RISTI debe hacerse de conformidad con las llamadas de los artículos y las instrucciones y normas establecidas en el sitio Web de la revista (<http://www.risti.xyz/>).

Todos los trabajos enviados son evaluados por un número de miembros del Consejo Científico de no menos de tres elementos.

En cada número de la revista se publican cinco a ocho de los mejores artículos enviados.

Os asociados da AISTI recebem a RISTI gratuitamente, por correo postal. Torne-se associado da AISTI. Preencha o formulário abaixo e envie-o para o e-mail aistic@gmail.com

Los asociados de la AISTI reciben la RISTI por correo, sin costo alguno. Hazte miembro de la AISTI. Rellena el siguiente formulario y remítelo al e-mail aistic@gmail.com



Formulário de Associado / Formulario de Asociado

Nome/Nombre: _____

Instituição/Institución: _____

Departamento: _____

Morada/Dirección: _____

Código Postal: _____ Localidade/Localidad: _____

País: _____

Telefone/Teléfono: _____

E-mail: _____ Web: _____

Tipo de Associado e valor da anuidade:

☐ Individual - 25€

☐ Instituição de Ensino ou I&D/Institución de Educación o I&D - 250€

☐ Outro (Empresa, etc.) - 500€

NIF/CIF: _____

Data/Fecha: __/__/____ Assinatura/Firma: _____



Revista Ibérica de Sistemas e Tecnologias de Informação
Revista Ibérica de Sistemas y Tecnologías de Información

©AISTI 2017 <http://www.aisti.eu>

