

La Simulación en Ingeniería

Trascendiendo Fronteras



Copyright: Dirección de Investigación de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas, prohibida la reproducción total o parcial de este libro por ningún medio impreso o electrónico sin el permiso previo y por escrito del dueño del copyright.

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR SEDE ESMERALDAS

Dirección: Espejo y subida a Santa Cruz
Casilla: 08-01-0065

Teléfonos: +593 (06) 2721983 – 2721595

Email: direccion.sistemas@pucese.edu.ec

www.pucese.edu.ec

ESMERALDAS - ECUADOR

PUBLICACIÓN ELECTRÓNICA

ISBN: 978-9942-9902-6-6

Los artículos incluidos en esta publicación fueron sometidos a procedimientos de admisión y revisión por pares, llevados a cabo por un comité científico nacional e internacional de alto nivel.

La simulación en ingeniería, trascendiendo fronteras

Coordinador de la Carrera de Sistemas y Computación

Víctor Xavier Quiñonez Ku

Director de Investigación PUCE Esmeraldas

Ignacio Carazo Ortega

Comité Organizador

Víctor Xavier Quiñonez Ku

Jaime Paúl Sayago Heredia

Luis Alberto Herrera Izquierdo

Comité Editorial

Pablo Antonio Pico Valencia, *Universidad de Granada – España*

Evelin Lorena Flores García, *Universidad de Almería – España*

Cesar Raúl García Jacas, *Universidad Nacional Autónoma de México*

José Luis Sampietro Saquicela, *Universidad de Barcelona – España*

Luis Dionicio Rosales Romero, *Universidad de los Andes – Venezuela*

Juan Luis Casierra Cavada, *Pontificia Universidad Católica del Ecuador*

Pedro Roberto Suarez Suri, *Pontificia Universidad Católica del Ecuador*

Víctor Xavier Quiñonez Ku, *Pontificia Universidad Católica del Ecuador*

Jaime Paúl Sayago Heredia, *Pontificia Universidad Católica del Ecuador*

Luis Alberto Herrera Izquierdo, *Pontificia Universidad Católica del Ecuador*

Franyelit María Suarez Carreño, *Pontificia Universidad Católica del Ecuador*

Marc Grob, *Pontificia Universidad Católica del Ecuador*

Comité Técnico

Manuel Rogelio Nevárez Toledo

Wilson Gustavo Chango Sailema

Susana Gabriela Patiño Rosado

José Luis Carvajal Carvajal

Pedro César Godoy Rosero

Junior Williams Mera Quiroz

Kleber Rolando Posligua Flores

Víctor Fabián Martínez Estupiñán

Walter Eduardo Almeida Clavijo

Verónica del Carmen Yáñez Ortiz

Ángel Anchundia Ortiz

Rhay Pablo Bustos Vera

Zully Betty Carvache Franco

Luis Enrique Hidalgo Solorzano

Walter Clemente Mosquera Torres

Tania Elizabeth Posligua Flores

Kleber Leonardo Vera Tortorella

Diseño Gráfico

Melba Cristina Marmolejo Cueva

Autor: PUCE Sede Esmeraldas

Editorial: Pontificia Universidad Católica del Ecuador Sede Esmeraldas

Colección: Investigación, Temas relacionados con la Tecnología

Fecha publicación: Enero 2018

Edición: 1^{ra} edición

Idioma: Español

PRESENTACIÓN



La Pontificia Universidad Católica del Ecuador Sede Esmeraldas (PUCESE) ha asumido como un reto la tarea de la investigación en el desarrollo académico. Fruto de ello han sido las I Jornadas de investigación de la carrera de Sistemas y Computación bajo el título *La simulación en Ingeniería: trascendiendo fronteras*.

La visión de la PUCESE ha sido siempre impulsar el desarrollo y mejorar la calidad de vida de las personas, especialmente las más vulnerables. La investigación en áreas técnicas también tiene este objetivo y eso es lo que se quiere poner de manifiesto en las presentes jornadas con la participación de profesionales, docentes y estudiantes de distintas universidades.

Además de poner en común los trabajos que vienen realizando investigadores en ingeniería, las jornadas quieren consolidar alianzas para seguir trabajando en redes de investigación que permitan avanzar cooperativamente de forma más eficiente e interdisciplinaria.

Estas jornadas se enmarcan también en un momento de cambio y rediseño de carreras, con los respectivos estudios de pertinencia para responder a las necesidades del país y las demandas de los estudiantes. En este sentido, dentro del área técnica, el próximo año 2018 tendremos una nueva propuesta académica, que se denomina Tecnologías de Información y que sustituye a la anterior de Sistemas y Computación. Un cambio importante en un área de conocimiento que evoluciona aceleradamente. Esperamos desde ella seguir aportando a la transformación de Esmeraldas.

Un cordial saludo

Aitor Urbina
PRO-RECTOR

ÍNDICE DE ARTÍCULOS

	PÁG.
MULTIMODAL HYPERDOCTRINES	5
GESTIÓN DE LUMINARIAS LED USANDO 6LOWPAN	12
IMPLEMENTACIÓN DE UN APLICATIVO PARA VERIFICAR LA CONECTIVIDAD DE NODOS DE SENSORES INALÁMBRICOS UTILIZANDO IPV6	17
SIMULADOR DE COGNITIVE RADIO EN REDES MÓVILES	23
SCRUM Y DJANGO PARA EL DESARROLLO RÁPIDO DE UN SISTEMA WEB PARA EL CONTROL DE MANTENIMIENTOS PREVENTIVOS Y CORRECTIVOS DE BIENES DEL CUERPO DE BOMBEROS DEL GADM DE SANTO DOMINGO	29
GENERADOR AUTOMÁTICO DE APLICACIONES WEB E INTERFACES DE USUARIOS CON FUNCIONALIDAD RESPONSIVA EN EL LENGUAJE PYTHON	37
ANÁLISIS DE SENTIMIENTO CON LINGMOTIF: APROXIMACIÓN METODOLÓGICA AL ANÁLISIS DE INTERVENCIONES EN SERIES TEMPORALES DE LA VALENCIA	43
INCIDENCIA DE LA COMUNICACIÓN VISUAL EN LA EVALUACIÓN HEURÍSTICA	52
DETERMINACIÓN DEL RITMO CARDÍACO EN TIEMPO REAL PARA EL DIAGNÓSTICO DE AFLICCIONES RELACIONADAS AL CORAZÓN MEDIANTE LA APLICACIÓN DE REDES NEURONALES	56
PARÁMETROS INCIDENTES EN EL CONSUMO ELÉCTRICO, REFERIDOS AL SECTOR RESIDENCIAL URBANO	61
LA VINCULACIÓN ENTRE LA UNIVERSIDAD Y EL SECTOR PRODUCTIVO INDUSTRIAL	66
MINERÍA DE OPINIONES DE UN PERSONAJE PÚBLICO EN EL ECUADOR	72
METODOLOGÍA PARA PROCESOS DE INTELIGENCIA DE NEGOCIOS CON MEJORAS EN EXTRACCIÓN Y TRANSFORMACIÓN DE FUENTES DE DATOS	77
ANÁLISIS DE LA RESPUESTA ESPECTRAL EN PLANTACIONES DE BANANO, A TRAVÉS DE LA RELACIÓN DE LOS MICRO ELEMENTOS DEL SUELO Y LOS ÍNDICES DE VEGETACIÓN, EN LA FINCA BLANCA AMÉRICA, PROVINCIA DEL GUAYAS, CANTÓN SAN JACINTO DE YAGUACHI	81
SISTEMAS DE CONTROL DISTRIBUIDO PARA EL PROCESO DE GENERACIÓN DE VAPOR EN UNA PLANTA TERMOELÉCTRICA	89
ANÁLISIS COMPARATIVO DE METODOLOGÍAS DE INTERPOLACIÓN DE POLINOMIOS QUE REPRESENTEN CONTORNOS DE MAPAS PARA CÁLCULO DE ÁREAS PLANAS	93

Multimodal hyperdoctrines

Joaquín Díaz Boils

Escuela de Ciencias Físicas y Matemáticas
 Facultad de Ciencias Exactas y Naturales
 Pontificia Universidad Católica del Ecuador
 170150. Quito. Ecuador.
 jdiaz987@puce.edu.ec

Abstract—This paper treats type systems addressed to the semantics of recursive function classes closed under the known as safe recursion scheme.

The strategy has been the introduction of an extension of former developments made for two sorts of variables which have no counterpart in the above-mentioned type systems. Since these typings make use of modal operators to perform safety, we need a suitable interpretation of multimodal systems.

The main goal has been achieved through the introduction of Multimodal Hyperdoctrines whose models, obtained through categorical interpretations of Boolean Algebras with Operators, should give an account of those semantics for recursion.

Keywords—*Fibrations, Grothendieck Construction, Hyperdoctrines, Multimodal Logic, Boolean Algebras with Operators.*

I. INTRODUCTION

Several different approaches are considered in this work. It contains an algebraic point of view by the consideration of Boolean Algebras with operators, a type-theoretic perspective and the use of logical systems all of them addressed to a better understanding of complexity, and some recursive classes in particular.

The systems considered to perform the idea expressed in the abstract contain a number of modal operators, following the ideas of [1], and that is precisely the starting point to introduce algebraic structures from which we can develop a categorical formulation of multimodal systems. Multimodal Hyperdoctrines appear in this context as an appropriate setting for these logic systems.

It turns out, then, that the categories arising from the Grothendieck Construction associated to these hyperdoctrines contain the representation of the models which we hope to use in the characterization of the complexity classes mentioned.

Section 2 introduces the basic ideas of safety in the context of computability. In section 3, following the development of [1] and the extension introduced by the author in [2], a modal type system, suitable for our purpose is presented as well as its meaning. Some example of recursive functions are given in that section. The key concept of Multimodal Hyperdoctrine is given in section 4 while in section 5 it is introduced a suitable monad for many modalities whose category of algebras is proved to be isomorphic to the category of Multi-Interior Algebras. In the last section 6 a sketch is given about which is the meaning of the categories arising from the Grothendieck Construction in this context.

II. SAFE RECURSION AND COMPOSITION

The categorical work on the complexity hierarchies contained in the Primitive Recursive Functions class has to deal with a characterization of the known as *bounded recursion*:

$$\begin{cases} f(u, 0) = g(u) \\ f(u, x + 1) = h(u, x, f(u, x)) \\ f(u, x) \leq j(u, x) \end{cases} \quad (1)$$

The *Grzegorzcyk Hierarchy* in particular is a well-known class which is closed under that scheme.

The safe interpretation was introduced by Bellantoni and Cook in [3] as a way to substitute the bounding condition in the above scheme by a syntactical condition. Their central idea was to define two different kinds of variables (*normal* and *safe*) according to the use we make of them in the process of computation. In [3] the class of polynomial time functions has been characterized in safety terms. In particular, Bellantoni and Cook define a class of functions in the form $f(\bar{x}, \bar{y})$ where normal inputs are in the left and separate them from safe by making use of a semicolon. From that seminal work two safe operations are defined in the following form:

- *Safe Recursion* corresponding to the scheme:

$$\begin{cases} f(0, \bar{x}; \bar{a}) = g(\bar{x}; \bar{a}) \\ f(y + 1, \bar{x}; \bar{a}) = h(y, \bar{x}; \bar{a}, f(y, \bar{x}; \bar{a})) \end{cases} \quad (2)$$

where g and h belong to the class considered and

- *Safe Composition* by means of which we can define a new function f by

$$f(\bar{x}; \bar{a}) = h(\bar{g}(\bar{x}); \bar{k}(\bar{x}; \bar{a})) \quad (3)$$

where h , \bar{g} and \bar{k} belong to the class considered.

The important point of safe composition is that safe inputs will never be moved to a normal position, this fact allows to control recursion in a way that is proved to be analogous to the bounding condition in the bounded recursive scheme referred to above.

A characterization of the Grzegorzcyk Hierarchy from safe operators is presented in [4] where generalized forms of safe

recursion and composition are introduced, the former having the form

$$\begin{cases} f(\bar{x}_n; \dots; \bar{x}_{k+2}; \bar{x}_{k+1}, 0_{k+1}; \bar{x}_k) = g(\bar{x}_n; \dots; \bar{x}_{k+2}; \bar{x}_{k+1}; \bar{x}_k) \\ f(\bar{x}_n; \dots; \bar{x}_{k+2}; \bar{x}_{k+1}, a_{k+1} + 1; \bar{x}_k) = \\ h(\bar{x}_n; \dots; \bar{x}_{k+2}; \bar{x}_{k+1}, a_{k+1}; \bar{x}_k, f(\bar{x}_n; \dots; \bar{x}_{k+2}; \bar{x}_{k+1}, a_{k+1}; \bar{x}_k)) \end{cases} \quad (4)$$

The proposals performing complexity in safety terms are based on explicit recursion theory or else on type theory, [3] is of the first kind since in the constraint is inside the syntax of the scheme. The second kind of approaches are purely based on type and deductive systems. Other systems, such as the one presented in the next section, can be classified between these two styles. That hybrid system present recursion in the context of a typed calculi by making use of other techniques (such as linearity) to control the computational growth ([1]).

III. A TYPED MULTIMODAL SYSTEM

We define a type system which captures the above generalized scheme by making use of a language containing a number of modalities in the sense of the (uni)modal system of [1]. In that reference an *affine closed category* is considered, a category with tensor product and projections closely related to cartesian closedness (as in [5]) and symmetric monoidal closedness (as in [6]) both of them endowed with a nno as shown in the quoted texts. We will not get into details about that, just to observe that linear function spaces can be considered in an affine closed category.

The typing introduced here contains both recursors and types which allow to interpret safe recursion in many variables. Our system, called SL R n as an extension of SLR in [1], have the features of a lambda calculus with many modalities and linear function spaces. In it, a number of cartesian-preserving operators is considered as well as the usual rules for grammar, typing and subtyping described in Appendix 1.

Our language contains different function spaces and a number of modalities performing the different sorted variables considered in the generalized safe recursion scheme (see section 2). The key point of this system is the ability of some function spaces (linear function spaces) to use their variables exactly once, this is why we consider both tensor and cartesian products to link the modalities. At the end we will be working with a weakening of Gödel's system T as done before (see [3,7]) by considering *linearity* as a way to introduce safety.

We begin by defining the basic types:

N	natural numbers
A	type variable
$\Box_i A$	modal type
$A \times B$	cartesian product
$A \otimes B$	tensor product
$A \multimap B$	linear function space
$\Box_{i+1} A \rightarrow \Box_i B$	i-function space
$A \rightarrow B$	function space

where $0 \leq i \leq n - 2$ and $\Box_0 A = A$.

Cartesian and tensor product bind stronger than function spaces and we associate to the right. These are the type expressions of a system we will call SLR_n .

The sense of having two different kinds of function space (\rightarrow and \multimap) has to do with the fact that we are using higher order types to define (safe) recursive operators. For them we need to define an ordering. The intuition behind the ordering above defined is exposed in [1], we will only point that our setting with an indeterminate number of modalities needs to observe more rules between the different function spaces considered.

We consider in the following a typing with n modalities and functions defined after at most $n-1$ recursions.

Definition 1. A k -aspect is a pair (l, \Box_k) where $l \in \{\rightarrow, \multimap\}$ and $k \in \{0, \dots, n - 1\}$ and where (\multimap, \Box_{n-1}) is not an aspect.

A k -aspect and a j -aspect are ordered componentwise by the rules

$$\rightarrow <: \multimap \quad \text{and} \quad \Box_i A <: \Box_k A \quad (5)$$

for $j \in \{k, \dots, n - 1\}$ where \multimap is said to be *linear* and \rightarrow *nonlinear*. That is a generic function space is of one of the following forms:

$$\begin{aligned} \Box_j A \multimap \Box_k B \quad \text{whenever } a = (\multimap, \Box_j) \quad \text{and } n - 2 \geq j \geq k \geq 0 \\ \Box_j A \rightarrow \Box_k B \quad \text{whenever } a = (\rightarrow, \Box_j) \quad \text{and } n - 1 \geq j \geq k \geq 0 \end{aligned}$$

subject to the following inequalities:

$$\Box_k A \multimap \Box_k B <: \Box_k A \rightarrow \Box_k B <: \Box_{k+1} A \rightarrow \Box_k B \quad (6)$$

where we reverse the order of aspects because of the rule

$$\frac{A' <: A \quad B <: B' \quad b <: a}{A \xrightarrow{a} B <: A' \xrightarrow{b} B'}$$

in table 2 of Appendix 1.

Variables typed $\Box_k A$ are values that have been computed in a previous level of the hierarchy than j for $k < j$, that is, they cannot be accessed in the current, higher complexity stage. The relation $<: B$, read *A is a subtype of B*, is defined inductively by the rules contained in Table 2 of Appendix 1 while $\Box_{k+1} A <: \Box_k A$ can be read, according to [1]: $\Box_{k+1} A$ offers more capabilities than $\Box_k A$ or $\Box_{k+1} A$ is more known than $\Box_k A$ in an impredicative interpretation. This is coherent with our many-modalities system in the sense that we are describing more complex functions.

A *context* is a partial function sending a term to a pair formed by an aspect and a type, we write them in the form $\Gamma \vdash f: A$. If Γ is a context we write $dom(\Gamma)$ for the set of variables bound in Γ and, if $x: A \in \Gamma$, then we write $\Gamma(x)$ for A . The typing relation between contexts, expressions and types is defined by the rules contained in Table 3 of Appendix 1. We suppose that if Γ, Δ appears as a premise or conclusion of a rule then Γ and Δ are disjoint. Finally, a context is linear/nonlinear if it is so in every variable of its domain.

Now we define constant terms for SLR_n :

$$\begin{aligned} 0_k &: \Box_k N \\ s_k &: \Box_k N \multimap \Box_k N \end{aligned}$$

and the recursion operators for level $k \leq n - 2$ in a subrecursive hierarchy:

- k -recursive operators for $k = 0, \dots, n - 2$

$$rec_k^N: \forall X. \Box_p N \multimap (X \rightarrow \Box_p N \multimap \Box_p N) \rightarrow X \rightarrow \Box_p N$$

where X has the typing

$$\Box_l N \rightarrow \dots \rightarrow \Box_k N$$

for $k, p \in \{0, \dots, n - 2\}$ and $l \in \{1, \dots, n - 1\}$ where $l > p$ for the variables typed $\Box_l A$ which are being currently recursed.

We are not using in these typings the nomenclature safe/normal anymore since it gives no clarity in an n-typed operator like rec_k^N .

We could also consider k -flat operators in the form

$$flat_k^N: \Box_k N \multimap (\Box_k N \multimap \Box_k N) \rightarrow \Box_k N \rightarrow \Box_k N$$

for $k = 0, \dots, n - 2$ to perform some important functions but we leave them to simplify the typing.

The typing here presented can be simplified: given a function in a system with n modalities obtained by recursion over a k -modal variable with $k < n$, we can obtain an *equivalent* function containing a $(k + 1)$ -modal input. At the same time, we can obtain an even simpler typing by making use of the rule which allows to lower the output level. Both features are given by rules such as

$$\frac{A' <: A \quad B <: B' \quad b <: a}{A \xrightarrow{a} B <: A' \xrightarrow{b} B'} \quad \frac{\Gamma \vdash f : A \quad A <: B}{\Gamma \vdash f : B}$$

contained in table 2 of Appendix 1.

As an example we consider a particular instance of Ackermann function with which we can express the growth of functions into the *Grzegorzcyk Hierarchy*. It was defined by Goodstein in [8], give an analogous description of the functions belonging to the so-called *Hyperoperation Sequence*, and represent an easy way to classify the functions into the mentioned hierarchy by its complexity.

Definition 2. Let $G: N^3 \rightarrow N$ the sequence of recursive functions defined by:

$$G(n, a, b) = \begin{cases} b + 1 & \text{for } n = 0 \\ a & \text{for } n = 1 \text{ and } b = 0 \\ 0 & \text{for } n = 2 \text{ and } b = 0 \\ 1 & \text{for } n \geq 3 \text{ and } b = 0 \\ G(n - 1, a, G(n, a, b - 1)) & \text{otherwise} \end{cases}$$

Every G grows faster than any function defined by less than n levels of nested primitive recursion. Following the

sequence for every $n > 2$ we see the reason for adopting the *Knuth's up-arrow notation* for G , that is:

$$G(n, a, b) = \uparrow^{n-2} \quad (7)$$

whose type is $(n - 1, n - 2; n - 3)$. *Addition, multiplication, exponentiation, tetration* and in general the n th-Goodstein function for $n > 3$ can be given in the form $G(n, -, -)$.

Example 1. We give the type of some known functions:

- *addition*

$$\Box_0 N \multimap ((\Box_1 N \rightarrow \Box_0 N) \rightarrow \Box_0 N \multimap \Box_0 N) \rightarrow (\Box_1 N \rightarrow \Box_0 N) \rightarrow \Box_0 N$$

- *exponentiation*

$$\Box_0 N \multimap ((\Box_2 N \rightarrow \Box_1 N) \rightarrow \Box_0 N \multimap \Box_0 N) \rightarrow (\Box_2 N \rightarrow \Box_1 N) \rightarrow \Box_0 N$$

- *Goodstein functions* $G(n, -, -)$

$$\Box_{n-3} N \multimap ((\Box_{n-1} N \rightarrow \Box_{n-2} N) \rightarrow \Box_{n-3} N \multimap \Box_{n-3} N) \rightarrow (\Box_{n-1} N \rightarrow \Box_{n-2} N) \rightarrow \Box_{n-3} N$$

IV. MODAL HYPERDOCTRINES

In order to have a consistent system with n modalities as described above, we consider the categorical concept of *hyperdoctrine*. Our choice is justified because, instead of the usual monad/comonad interpretation (see [9,10]), we find clearer its logical interpretation.

We are going to generalize the concept of *Interior Algebra* to a many modalities concept. With it we will introduce the concept of *Multimodal Hyperdoctrine* which encompass the main features of the multimodal systems we want to characterize categorically. An interior algebra is a kind of *Modal Algebra* which in its turn is a Boolean Algebra (BA in the sequel) with one operator, the simplest case of Boolean Algebra with Operators (BAO in the sequel).

Definition 2. An *Interior Algebra* B is a $BA (B, \wedge, \vee, \top, \perp, \neg)$ with an operator \Box such that

$$\Box(a \wedge b) = \Box a \wedge \Box b \quad \Box \top = \top \quad \Box \Box a = \Box a \quad \Box a \leq a$$

The last axiom means $\Box a \wedge a = \Box a$.

There is a 2-category *Int* whose objects are interior algebras. The concept of Interior Algebra can be extended by considering a set of operators, this means taking a special instance of BAO .

Definition 3. A *Multi-interior Algebra* B_n is a $BA (B, \wedge, \vee, \top, \perp, \neg)$ with a set of n operators \Box_i for $i = 1, \dots, n$ such that

$$\Box_i(a \wedge b) = \Box_i a \wedge \Box_i b \quad \Box_i \top = \top \quad \Box_i \Box_i a = \Box_i a \quad \Box_i a \leq a$$

We denote by $(B, \Box_1, \dots, \Box_n)$ the *multi-interior algebras* which are the objects of an obvious 2-category $MInt_n$.

Arrows $h: B \rightarrow B'$ between $(B, \square_1, \dots, \square_n)$ and $(B, \ominus_1, \dots, \ominus_n)$ in that category are homomorphisms between BAs satisfying the following coherent condition

$$h(\square_i a) = \ominus_i h(a) \quad (8)$$

for every $a \in B$.

As in the case of Int , for which we can define a *Modal Hyperdoctrine* (see [11]), we define from $MInt_n$ a *Multimodal Hyperdoctrine*. This is done in the following.

Definition 4. Given a cartesian category \mathcal{C} a *Multimodal Hyperdoctrine* is a functor

$$M: \mathcal{C}^{op} \rightarrow MInt_n$$

such that:

1. for every $f: Y \rightarrow X$ in \mathcal{C} the reindexing functor $M(f): M(X) \rightarrow M(Y)$ has a left adjoint $\exists(f): \exists(Y) \rightarrow \exists(X)$ and a right adjoint $\forall(f): \forall(Y) \rightarrow \forall(X)$
2. given a pullback

$$\begin{array}{ccc} Y \times Z & \xrightarrow{f \times Z} & X \times Z \\ \pi_0 \downarrow & & \downarrow \pi_0 \\ Y & \xrightarrow{f} & X \end{array}$$

then for every $A \in M(X \times Z)$ the following diagrams commute

$$\begin{array}{ccc} M(X \times Z) & \xrightarrow{\exists(\pi_0)} & M(X) \\ M(f \times Z) \downarrow & & \downarrow M(f) \\ M(Y \times Z) & \xrightarrow{\exists(\pi_0)} & M(Y) \end{array} \quad \begin{array}{ccc} M(X \times Z) & \xrightarrow{\forall(\pi_0)} & M(X) \\ M(f \times Z) \downarrow & & \downarrow M(f) \\ M(Y \times Z) & \xrightarrow{\forall(\pi_0)} & M(Y) \end{array}$$

Condition 2. is known as the *Beck-Chevalley condition*. We recall that the objects $M(X)$ are posets according to definition 4, this fact allows to interpret the adjunctions as inequalities

$$\begin{array}{ll} B \leq M(f)(A) \rightarrow \exists(f)(B) \leq A & B \leq M(f)\exists(f)(B) \\ M(f)(A) \leq B \rightarrow A \leq \forall(f)(B) & M(f)(B)\forall(f) \leq B \end{array}$$

for every $A \in M(Y), B \in M(X)$.

The adjunctions above are known to characterize the existence and universal quantifications hence the notation (see [13]). We use *reindexing functor* rather than *inverse image functor* because we want to focus in the fibrational interpretation. On the other hand, we will not make use of the Lindenbaum-Tarski interpretation of modal hyperdoctrines in terms of natural numbers. This syntactical definition should be the motivation to introduce models which give an appropriate interpretation of concrete multimodal systems. This could be considered as future work for this paper.

V. ALGEBRAS FOR A FUNCTOR OF COPRODUCTS

We interpret our categories $MInt_n$ as (isomorphic to) algebras for a certain monad. This will help to understand

which is the form of their objects in terms of Boolean Algebras (essentially, a coproduct of BAs). Following [14] we consider a pair of adjoint functors between the categories of BAs and the category of semilattices which give rise to a monad whose category of algebras is isomorphic to $MInt_n$, this has been suggested in [14] as a way to express the internal structure of BAOs.

Let be the adjunction

$$\begin{array}{ccc} & U & \\ \text{Bool} & \xrightarrow{\quad} & \text{SemiLat} \\ & P & \\ & \perp & \end{array}$$

from which we can construct a monad

$$PU: \text{Bool} \rightarrow \text{Bool}$$

which we will denote T_n whenever the BA considered is exactly a BAO with n operators. They are precisely the algebras of monads in this form, which we will describe in detail below, that we claim to be isomorphic to $MInt_n$. From that fact we will give an analogous definition of hyperdoctrine in the form

$$\mathcal{C}^{op} \rightarrow \text{Alg}(T_n)$$

In particular:

Theorem 1. There is an isomorphism functor

$$\Psi: MInt \rightarrow \text{Alg}(T_n)$$

Proof. See Appendix 2.

The adjunction above allows to consider bijections

$$\text{SemiLat}[R, U(B)] \xrightarrow{(-)^{T_n}} \text{Bool}[P(R), B]$$

$$\text{Bool}[P(R), B] \xrightarrow{(-)^{T_n^{-1}}} \text{SemiLat}[R, U(B)]$$

for R a semilattice and B a BA. Now, for $\psi: R \rightarrow U(B)$, $\varphi: S \rightarrow R \in \text{SemiL}$ and $\mu: B \rightarrow C \in \text{Bool}$, we have the following equations obtained by adjointness

$$\begin{array}{l} (U(\mu)\psi)^{T_n} = \mu\psi^{T_n} \\ (\psi\varphi)^{T_n} = \psi^{T_n}P(\varphi) \end{array}$$

and analogous equations for the functor $(-)^{T_n^{-1}}$.

Definition 5. Given a monad T over a category \mathcal{C} the *category of algebras over T* denoted by $\text{Alg}(T)$ has:

- as objects pairs $[C, f: TC \rightarrow C]$
- as morphisms between the objects $[C, f: TC \rightarrow C]$ and $[C', f': TC' \rightarrow C']$ in $\text{Alg}(T)$ a morphism h in \mathcal{C} such that the following diagram commutes

$$\begin{array}{ccc}
TC & \xrightarrow{T_h} & TC' \\
f \downarrow & & \downarrow f' \\
C & \xrightarrow{h} & C'
\end{array}$$

Let $(B, \square_1, \dots, \square_n)$ be a Multi-interior algebra, from it we can construct an algebra for T_n in the form $(B, \square_1^{T_n}, \dots, \square_n^{T_n})$, it is exactly an object $[B, T_n(B) \rightarrow B] \in Alg(T_n)$. What we need to know is in which form T_n acts over an arbitrary BAO . The form in which our monad T_n acts has been suggested in [14]. It expresses every BAO with n operators in the form of a coproduct of n BAO s with just one operator. This fact can be seen categorically (the one we choose for this article) or algebraically in terms either of lattices or modules. In particular, every BA is a lattice so we could construct the coproduct of lattices but a BA is at the same time a module over the Galois ring of two elements, in this second case it turns out that the tensor product of these modules is also a coproduct in the categorical sense. In any of the cases we are obtaining finite joins of finite meets of elements of the disjoint union of BA s as the elements of a coproduct of those BA s (see [15]).

This appears to be perhaps the easiest way to express a coproduct of BAO s and the one we will consider. With this fact in mind we observe that our functor T_n acts sending a BA to a coproduct of n copies of that BA every one of them endowed with one of the operators considered. That is,

$$T_n(B) = \coprod_{i=1}^n P(U(B))_i = P(U(B))_1 \oplus \dots \oplus P(U(B))_n \quad (9)$$

where every $P(U(B))_i$ is a BAO with a single operator \square_i . We are relying for these matters on the facts that the category $Bool$ has coproducts and that they are preserved by the functors U and P .

VI. FIBRATIONAL BOOLEAN ALGEBRA WITH OPERATORS

We are getting into the description of a framework to describe the semantics of a multimodal formalism in an indexed manner. This *indexed semantic*, quite usual in Universal Algebra and Categorical Logic, makes use of the category Set as the chosen semantic environment. For that, we will show what the Grothendieck Construction of our multimodal hyperdoctrines have to do with concrete models of the specification considered.

Since a hyperdoctrine $M: \mathcal{C}^{op} \rightarrow MInt_n$ can be seen as an indexed category we can construct a cloven fibration (see [12])

$$\int_{\mathcal{C}} M \rightarrow \mathcal{C}$$

where the category $\int_{\mathcal{C}} M$, obtained through the known as *Grothendieck Construction*, has

- as objects pairs (C, X) where $C \in \mathcal{C}$ and $X \in M(C)$
- as arrows from (C, X) to (D, Y) pairs (φ, ϕ) where $\varphi: C \rightarrow D \in \mathcal{C}$ and $\phi: X \rightarrow Y \in M(C)$.

The categories $\int_{\mathcal{C}} M$ in indexed contexts as ours give rise to interpretations for a particular algebraic specification. Our claim is that, whenever take $\mathcal{C} = Set$, the multimodal models we are looking for to perform the complexity classes can be obtained through these categories $\int_{\mathcal{C}} M$. It is precisely in that case, whenever we work with the category of sets, that we have not just quantifiers for the hyperdoctrine but a whole structure for higher order logic. We denote in the sequel simply $\int M$ for $\int_{Set} M$.

Definition 6. Given a multimodal language \mathcal{L} (in the sense of [16]) a *multimodal hyperdoctrinal model* is a tuple (M, B, Θ) where M is a multimodal hyperdoctrine, B a set and Θ an interpretation

$$\Theta: \mathcal{L} \rightarrow Set$$

such that $\Theta(p)$ belongs to $M(B)$ for p a predicate in the language \mathcal{L} and Θ satisfies:

$$\begin{aligned}
\Theta(p \wedge q) &= \Theta p \wedge \Theta q & \Theta(\square_i p) &= \square_i \Theta(p) & \Theta(p \rightarrow q) &= \Theta(p) \rightarrow \Theta(q) \\
\Theta(\top) &= \top & \Theta(\perp) &= \perp
\end{aligned}$$

for p, q in \mathcal{L} as well as the interaction with the equality quantifier

$$\Theta(=) = \exists_{\Delta}(\top) \quad (10)$$

where Δ is a diagonal arrow.

The category obtained through the Grothendieck Construction has as objects pairs whose second coordinate consist of models for a specification of a certain axiomatic. In particular, the category $\int M$ for $M: Set^{op} \rightarrow MInt_n$ has as objects pairs (B, X) where $B \in Set$ and $X \in M(B)$ is a Multi-Interior Algebra with n operators. We see our algebras as models for the varieties generated by a particular specification (we think in this paper on the axiomatic of multimodal logics as in [17]). The Grothendieck Construction on an indexed category form is in our setting the context to interpret models or algebras for sound logics.

VII. CONCLUSIONS

It has been presented a multimodal type system suitable for representing recursive functions into complexity hierarchies. In particular, those closed under safe recursive operators performing bounded recursion schemes.

Subsequently, it has been given categorical interpretations of the systems introduced both considering Boolean Algebras with operators (BAO) and Hyperdoctrines. A coproduct of BAOs is then considered to perform the several modalities involved.

Finally, it has been given a survey of how the Grothendieck construction can be applied to the structures defined so far in a fibrational fashion.

With all this machinery at hand, we hope to have at our disposal a clearer presentation of multimodal systems in terms of the language of Category Theory.

REFERENCES

- [1] M. Hofmann. Typed lambda calculi for polynomial-time computation. Habilitation Thesis, TU Darmstadt, Germany. Edinburgh University LFCs Technical Report, ECS-LFCs-99-406, 1999.
- [2] Joaquín Díaz. *Categorical Comprehensions and Recursion*. Journal Logic Computation (2016) doi: 10.1093/logcom/exw020.
- [3] S. Bellantoni, S. Cook. New recursion-theoretic characterization of the polytime functions. *Comput. Complexity* 2, 1992, pp. 97 - 110.
- [4] Marc Wirz. Characterizing the Grzegorzczk hierarchy by safe recursion. Retrieved on 2009-04-21, 1999.
- [5] J. Lambek, P.J. Scott. Introduction to Higher Order Categorical Logic, CUP, 1986.
- [6] I. Mackie, L. Román, and S. Abramsky, An internal language for autonomous categories, *Journal of Applied Categorical Structures*, 1, 311–343, 1993.
- [7] Stephen J. Bellantoni, Karl-Heinz Niggl, Helmut Schwichtenberg. Higher type recursion, ramification and polynomial time. *Annals of Pure and Applied Logic*. Volume 104, Issues 13, 15 July 2000, Pages 1730.
- [8] R. L. Goodstein. *Transfinite Ordinals in Recursive Number Theory*. Journal of Symbolic Logic 12 (4): 123-129. 1947.
- [9] G Bierman, V de Paiva, *On an intuitionistic modal logic*, *Studia Logica* 65 (3), 383-416. 2000.
- [10] S. Kobayashi. *Monad as modality*. Theoretical Computer Science archive Volume 175 Issue 1, March 30, 1997 Pages 29 - 74.
- [11] Handbook of Modal Logic edited by Patrick Blackburn, Johan F.A.K. van Benthem, Frank Wolter. 2006.
- [12] B. Jacobs, *Categorical Logic and Type Theory*, Studies in Logic and the Foundations of Mathematics 141, North Holland, Elsevier, 1999. ISBN 0-444-50170-3.
- [13] Bill Lawvere, Equality in hyperdoctrines and comprehension schema as an adjoint functor, Proceedings of the AMS Symposium on Pure Mathematics XVII (1970), 1-14.
- [14] Ghilardi, S. (1995). An algebraic theory of normal forms. *Annals of Pure and Applied Logic*, 71(3), 189-245.
- [15] Givant, Steven; Halmos, Paul (2009). Introduction to Boolean Algebras. Undergraduate Texts in Mathematics, Springer. ISBN 978-0-387-40293-2.
- [16] D. M. Gabbay, A. Kurucz, E Wolter, and M. Zakharyashev. *Many-dimensional modal logics: theory and applications*. Studies in Logic and the Foundations of Mathematics, vol. 148. Elsevier, Amsterdam, xiv + 747. 2003.
- [17] Robert Goldblatt. *Algebraic Polymodal Logic: A Survey*, LOGIC JOURNAL OF THE IGPL, 2000, volume 8, number 4, pages 393-450.

APPENDIX 1

Grammar:		
(e ₁ e ₂)		application
λx : A.e		abstraction
AX.e		type abstraction
e[A]		type application
< e ₁ , e ₂ >		pairing for ×
e.1		first projection
e.2		second projection
e ₁ ⊗ e ₂		pairing for ⊗
let e ₁ = x ⊗ y in e ₂		⊗ elimination
let e ₁ = x in y		abbreviation
c		constant

Table 1.

Rules for subtyping:

$$\begin{array}{c}
 A <: A \\
 \frac{A' <: A \quad B <: B' \quad b <: a}{A \xrightarrow{a} B <: A' \xrightarrow{b} B'} \quad \frac{A <: B}{\Box_j N \rightarrow A <: \Box_j N \rightarrow B} \\
 \frac{A <: A' \quad B <: B'}{A \times B <: A' \times B'} \quad \frac{A <: A' \quad B <: B'}{A \otimes B <: A' \otimes B'}
 \end{array}$$

Table 2.

Rules for typing $\Gamma \vdash f : A$:

$$\begin{array}{c}
 \frac{x \in \text{dom}(\Gamma)}{\Gamma \vdash x : \Gamma(x)} \quad \frac{c : A}{\Gamma \vdash c : A} \quad \frac{\Gamma \vdash f : A \quad A <: B}{\Gamma \vdash f : B} \\
 \frac{\Gamma, x : A \vdash f : B}{\Gamma \vdash \lambda x : A. f : A \rightarrow B} \quad \frac{\Gamma, \Delta_1 \vdash f_1 : A \xrightarrow{a} B \quad \Gamma, \Delta_2 \vdash f_2 : A \quad \Gamma \text{ nonlinear} \quad \Gamma, \Delta_2 <: a}{\Gamma, \Delta_1, \Delta_2 \vdash (f_1 f_2) : B} \\
 \frac{\Gamma \vdash f : \forall X. A \quad S \text{ safe}}{\Gamma \vdash f[S] : A[S/X]} \quad \frac{\Gamma \vdash f : A \quad X \text{ not free in } \Gamma}{\Gamma \vdash \lambda X. f : \forall X. A} \\
 \frac{\Gamma \vdash f_1 : A_1 \quad \Gamma \vdash f_2 : A_2}{\Gamma \vdash (f_1, f_2) : A_1 \times A_2} \quad \frac{\Gamma \vdash f : A_1 \times A_2 \quad i \in \{1, 2\}}{\Gamma \vdash f.i : A_i} \\
 \frac{\Gamma, \Delta_1 \vdash f_1 : A_1 \quad \Gamma, \Delta_2 \vdash f_2 : A \quad \Gamma \text{ nonlinear}}{\Gamma, \Delta_1, \Delta_2 \vdash f_1 \otimes f_2 : A_1 \otimes A_2} \quad \frac{\Gamma \vdash f : A \quad A <: B}{\Gamma \vdash f.i : A_i} \\
 \frac{\Gamma, \Delta_1 \vdash f_1 : A_1 \quad \Gamma, \Delta_2, x : A_1 \vdash f_2 : B \quad \Gamma \text{ nonlinear} \quad \Delta_1 \leq a_1}{\Gamma, \Delta_1, \Delta_2 \vdash \text{let } f_1 = x \text{ in } f_2 : B} \\
 \frac{\Gamma, \Delta_1 \vdash f_1 : A_1 \otimes A_2 \quad \Gamma, \Delta_2, x : A_1, y : A_2 \vdash f_2 : B \quad \Gamma \text{ nonlinear} \quad \Delta_2 \leq a_1 \wedge a_2}{\Gamma, \Delta_1, \Delta_2 \vdash \text{let } f_1 = x \otimes y \text{ in } f_2 : B}
 \end{array}$$

Table 3.

APPENDIX 2

Proof. of Theorem 1. Define

$$\Psi : MInt_n \rightarrow Alg(T_n)$$

as

$$\Psi(B, \Box_1, \dots, \Box_n) = (B, \Box_1^{T_n}, \dots, \Box_n^{T_n})$$

Then $f : (B, \Box_1, \dots, \Box_n) \rightarrow (B', \Theta_1, \dots, \Theta_n)$ is an arrow in $MInt_n$ precisely when it preserves the modal operators, that is:

$$\Box_i(U(f)(a)) = U(f)(\Theta_i(a'))$$

for every $a \in B, a' \in B'$ which means exactly

$$(\Box_i(U(f)(a)))^{T_n} = (U(f)(\Theta_i(a')))^{T_n}$$

if and only if

$$\Box_i^{T_n} f = P(U(f))(\Theta_i)^{T_n}$$

This last equation gives the commutative square

$$\begin{array}{ccc}
 \prod_{i=1}^n P(U(B))_i & \longrightarrow & \prod_{i=1}^n P(U(B'))_i \\
 \downarrow & & \downarrow \\
 B & \longrightarrow & B'
 \end{array}$$

which means that f is an arrow in $Alg(T_n)$ by taking $R = U(B)$ in the equations obtained by the adjunctions. Recall in all these equalities the adjunction

$$\begin{array}{ccc} & U & \\ \text{Bool} & \xrightarrow{\quad} & \text{SemiLat} \\ & P & \end{array}$$

and the assignments

$$U(f): \text{SemiLat}(B) \rightarrow \text{SemiLat}(B') \quad \Box_i: \text{SemiLat}(B) \rightarrow \text{SemiLat}(B)$$

$$\ominus_i: \text{SemiLat}(B') \rightarrow \text{SemiLat}(B')$$

because we need the modal operators preserve just the semilattice structure.

Gestión de luminarias led usando 6lowpan

Egas Acosta Carlos¹, Arellano Patricio², Casierra, Juan³
 cregas@puce.edu.ec, duck_arellano@hotmail.com, casierrajuan@gmail.com

¹Pontificia Universidad Católica del Ecuador- Ecuador

²Universidad de las Américas - Ecuador

³Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador

Resumen— La gestión de luminarias con el propósito de optimizar el uso de energía, representa un costo elevado cuando el sistema de transmisión de datos para control de las luminarias, requiere un sistema de comunicaciones cableado para interconectar los nodos sensores ubicados en cada luminaria. Este tipo de aplicaciones se caracteriza por los miles de luminarias que requieren ser gestionadas en forma remota, para lo cual es necesario la utilización de gran cantidad de direcciones IP públicas lo cual es un problema cuando se utiliza IPv4, además los nodos sensores requieren de hardware que les permita tomar energía de la red de eléctrica, lo cual encarece al nodo sensor. Algunas técnicas inalámbricas se han utilizado para interconectar los nodos sensores ubicados en las luminarias, muchas de esas técnicas utilizan innecesariamente anchos de banda grandes en relación a la cantidad de datos que se requieren transmitir resultado del sensado de las luminarias. Para resolver estos inconvenientes, proponemos la utilización de una red de sensores inalámbricos usando 6LoWPAN para la gestión de luminarias LED, el prototipo implementado para evidenciar la validez de nuestra propuesta utiliza nodos sensores de Libelium con el firmware de IBM.

Palabras clave: WSN, 6LoWPAN, Libelium, Mote Runner, IEEE 802.15.4

I. INTRODUCCIÓN

Las redes de sensores inalámbricos se basan en dispositivos conocidos como nodos de bajo costo y bajo consumo de energía, capaces de obtener información del lugar donde han sido ubicados, procesarla localmente y enviar los datos generados por el sensado a un nodo central conocido como Gateway utilizando enlaces inalámbricos. Estas redes están diseñadas para soportar miles de nodos en una sola zona de cobertura inalámbrica y si la zona a cubrir es extensa se pueden tener redes multisalto con nodos sensores fijos o móviles.

El estándar creado por el Instituto de Ingenieros Eléctricos y Electrónicos, conocido como IEEE (Institute of Electrical and Electronics Engineers), denominado IEEE 802.15.4, es la base para la operación de las redes sensores inalámbricas WSN (Wireless Sensor Networks) [1] sobre la cual se han desarrollado protocolos tales como ZigBee[2] y 6LoWPAN[3], los nodos son diseñados de tal manera que consuman la mínima cantidad de energía, ya que fueron diseñados para trabajar con baterías, es por esta razón que el estándar 802.15.4 define un estado de operación denominado modo sleep, en el cual el nodo tiene un mínimo consumo de energía ya que no están habilitados los procesos de transmisión y recepción.

Las WSN tienen capacidad de auto-restauración, por la cual, si un nodo no está operativo, y por lo tanto hay un cambio de topología, la red encuentra un nuevo camino

para encaminar los paquetes de datos. De esta forma, a pesar de que los nodos dejen de formar parte de la topología, la red continuará funcionando. A diferencia de otras tecnologías, las WSN presentan capacidades de auto-diagnóstico, auto-configuración, auto-organización, auto-restauración y reparación.

ZigBee es un grupo industrial formado para establecer las especificaciones que permitan aplicaciones inalámbricas fiables, económicas y de bajo consumo de energía basadas en la norma IEEE 802.15.4. y es la más difundida. Los desarrolladores iniciales de ZigBee está conformado por seis promotores principales (Honeywell, Samsung, Mitsubishi, Invensys, Motorola y Philips) y más de 80 participantes [2].

Debido a que el estándar 802.15.4 define una unidad de datos de 127 bytes, no se puede garantizar la transmisión de paquetes IP. Para que esto sea posible se debe comprimir las cabeceras, como lo propone el estándar 6LowPan el cual permite que los paquetes IPv6 se puedan transmitir utilizando redes que implementen el IEEE 802.15.4 [4]

En Octubre del 2013 fruto de una alianza entre IBM (*International Business Machines*) y Libelium [5], uno de los principales proveedores de hardware para redes sensores inalámbricas, lanzaron al mercado el kit de desarrollo para IoT denominado “Waspmote Moterunner” [6]. Los autores de esta investigación implementan el prototipo de la red WSN 6LowPan, en la cual se verifico la validez del Sistema de Gestión de Luminarias desarrollado.

II. 6LOWPAN

El estándar 6LoWPAN apareció con el objetivo de permitir el transporte de paquetes IPv6 sobre tramas IEEE 802.15.4, [7] permite acceder a los nodos utilizando direcciones ipv6, dispone de cientos de direcciones públicas para los nodos sensores necesarias para su gestión remota y los nodos sensores pueden conectarse directamente a la red internet sin necesidad de gateways o proxies, es decir estableciendo conexiones de extremo a extremo lo cual facilita al gestión.

A. Arquitectura

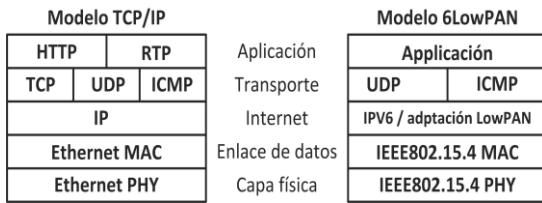


Fig. N°1. Modelos TCP/IP y 6LoWPAN

En la figura N° 1 se observa la comparación en los modelos TCP/IP y 6LoWPAN, la capa de adaptación LoWPAN es la encargada de adaptar IPv6 en 802.15.4, por ejemplo, para comprimir las cabeceras.

En 6LoWPAN, los nodos de la red se pueden auto configurar, este proceso se ejecuta en la red cuanto los nodos entran o salen de la misma y es realizado por la capa de enlace, lo que permite la conexión entre los nodos de la red, además, utiliza una versión modificada del protocolo de descubrimiento de vecinos ND (Neighbor Discovery) de IPv6.

B. Bandas de comunicación

El estándar IEEE 802.15.4 con el que trabaja 6LoPAN permite el uso de tres bandas de comunicación como se indica en la figura N°2, 868MHz (20Kbps), 915MHz (40Kbps) y 2.4GHz (250Kbps). La velocidad de transmisión típica es menor a 20Kbps.

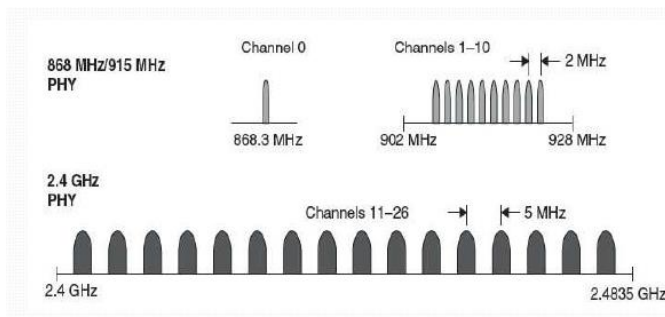


Fig. N°2. Bandas de frecuencia en IEEE 802.15.4

C. Estructura del paquete.

El formato del paquete de 6LoWPAN se muestra en la figura N° 3. Puesto que el estándar encapsula paquetes IPv6 sobre tramas de enlace IEEE 802.15.4, se define la capa de adaptación que se encarga de realizar el proceso de permitir trabajar con IPv6 en la arquitectura de una red WSN, añadir cabeceras adicionales y también ofrece mecanismos como fragmentación o compresión de cabeceras de capas superiores.

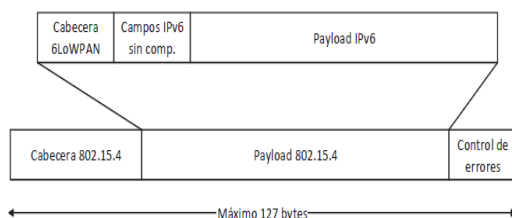


Fig. N° 3. Formato del paquete 6LoWPAN

Las cabeceras de adaptación inician con un campo de 8 bits conocido como dispatch byte, el mismo que sirve para identificar el tipo de cabecera, tal como se muestra en la figura N° 4

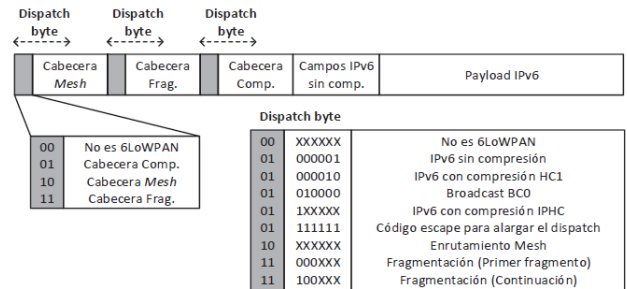


Fig. N° 4. Cabeceras de adaptación 6LoWPAN

III. HERRAMIENTA DE DESARROLLO

El kit de desarrollo de software denominado SDK (software development kit) fue el Mote Runner [8], desarrollada por los laboratorios de investigación de IBM en Zurich, la cual es una plataforma de desarrollo para redes sensores inalámbricas (WSN), que incluye herramientas de simulación de los nodos, herramientas para la configurar y grabar los programas en los nodos .

Los nodos sensores denominados waspmote, trabajan con el firmware Mote Runner, el cual ofrece un entorno de tiempo de ejecución, que incorpora una máquina virtual (VM) para la ejecución del sistema operativo (OS) y el micro código que permiten programar las diversas aplicaciones a ser ejecutadas en el waspmote de Libelium y otros fabricantes. La máquina virtual (VM) permite un único hilo de ejecución. Sin embargo, varias aplicaciones pueden ejecutarse simultáneamente en la VM.

El SDK cuenta con un conjunto de herramientas tales como interfaz de línea de comandos, aplicaciones basadas en web y servidores web para administrar y visualizar los waspmote de la WSN, ejecutando el sistema operativo Mote Runner y la máquina virtual VM.

Las aplicaciones en el waspmote pueden ser desarrolladas en C# y Java. Estas aplicaciones son compiladas, descargadas, instaladas, ejecutadas y depuradas en el simulador que se ejecuta en el computador. El Mote Runner proporciona las librerías para trabajar con 6LoWPAN y los waspmote de Libelium y únicamente están implementadas en C#, estas librerías deben ser importadas en la herramienta de desarrollo de C# a utilizar.

El firmware que habilita a las motas para ser utilizadas, es instalado en los waspmotes con ayuda del programa AVRdude o AVR Studio. Una vez cargado el firmware en cada waspmote, con los comandos del Mote Runner se

carga los archivos que contienen las aplicaciones para que puedan ser ejecutadas[9].

A. Kit de desarrollo waspmote

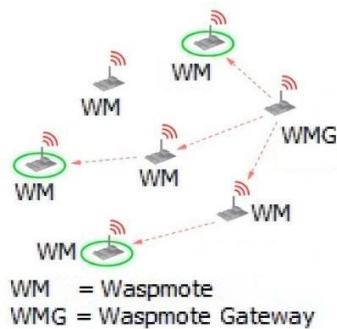


Fig. N° 5. Topología con los nodos waspmote

El kit utilizado, dispone de los siguientes tipos de nodos sensores que se indican en la figura N° 5.

Gateway: Toma la información enviada por los nodos (paquetes 6LoWPAN) y transforma las cabeceras IP a IPv4 manteniendo la capa de transporte UDP. Se encarga también de mantener una base de datos con el registro de los nodos conectados de forma que se reduzca el tráfico dentro de la red 6LoWPAN. Libelium denomina a estos nodos como Waspote Gateway (WMG)

Nodos sensores; Tienen integrados los sensores y se los conocen como Waspote (WM) dentro de la implementación de Libelium. Se encargan de obtener la información requerida y enviarla hacia el nodo WMG.

IV. SISTEMA DE GESTION DE LUMINARIAS

En la ciudad de Quito la gestión de alumbrado público está bajo responsabilidad de la Empresa Eléctrica Quito (EEQ). A febrero de 2016 existen 251.384 luminarias instaladas siendo una parte muy pequeña de LED. El monitoreo de estas luminarias es realizada con 21 grupos de trabajo que constan de tres personas cada uno utilizando diez vehículos especiales que se conocen como “canastillas”. Estos grupos trabajan en dos turnos de 07h00 a 15h00 y de 14h00 a 22h00 y se dividen en tres zonas para cubrir el Distrito Metropolitano. El costo de monitoreo utilizando estos grupos de trabajo bordea los \$756000 en pago al personal y \$120000 en el mantenimiento de los vehículos [10].

Frente a esta realidad se propone implementar la red de sensores para la gestión de las luminarias, con el propósito de reducir los costos que implican realizar el monitoreo de las luminarias led, mejorando el tiempo de respuesta de reposición.

La validez de la propuesta, se la realizó, implementando el sistema desarrollado, en un prototipo de red sensor inalámbrica, que permita gestionar las luminarias LED, utilizando sensores on-off para determinar si las luminarias

se encuentran encendidas y manejar estadísticas y conocer el estado de las luminarias en tiempo real, es decir encendida o apagada, el total del tiempo de uso y tiempo de duración de una luminaria permitiendo en este último caso la sustitución de la luminaria antes de que deje de operar.

A. Topología de la red

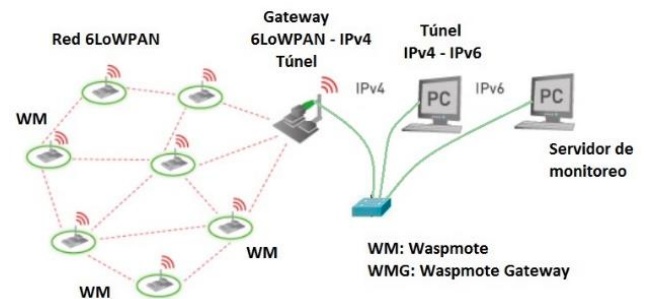


Fig. N° 6. Topología de red del prototipo

La el prototipo propuesto para la implementación del sistema de gestión se muestra en la topología de la figura N° 6. Se trata de una red de sensores inalámbricos con nodos sensores WM que estarían conectados a las luminarias LED que se desea gestionar. Los nodos utilizan 6LoWPAN y son controlados por medio del Gateway WMG el mismo que cambia la cabecera IP a IPv4 y mantiene la capa de transporte UDP. El prototipo en el cual se implementara el sistema de gestión de luminarias está configurado para transmitir a una velocidad de 250 kbit/s, utiliza CSMA/CD para la controlar el acceso al canal y los nodos son capaces de implementar funciones de gestión de energía a través de la detección de la calidad del enlace y la energía utilizada. Los nodos sensores envían la información a un servidor conectado a la red IPv6, el mismo que se encarga de almacenar y procesar los datos recolectados.

Para la conexión del Gateway con la PC, que en la práctica debe ser un solo equipo, se implementa un túnel IPv4/IPv6, esta es una condición del fabricante del kit de desarrollo. La PC, la cual estaría conectada al internet con IPv6, envía los datos al servidor, el cual también debe estar conectado a internet con IPv6, en el cual ejecuta la aplicación encargada de procesar los datos obtenidos. En este tipo de redes si falla el nodo WMG, los nodos sensores perderían comunicación hacia la red IPv6 y por tanto el sistema de gestión no recibiría los datos desde los mismos. Para lo cual se propone utilizar un nodo WMG de respaldo.

A. Software de comunicaciones.

Para poder procesar los datos obtenidos por los nodos es necesario que los mismos puedan llegar al servidor. La forma de realizar esto es implementando un software de comunicaciones con sockets tanto en el servidor de monitoreo conectado a internet con IPv6, así como en los nodos sensores WM en la WSN, de esta manera tener una

conectividad de extremo a extremo. Los sockets en los nodos WM tienen acceso directo a la capa de enlace. El software que se encuentra ejecutándose en los nodos WM, se encarga, entre otras actividades, de crear el paquete IPv6 que será transmitido por el socket en modo half dúplex.

En vista de que la programación para los nodos WM debe hacerse usando un hilo único de programación, los tiempos de procesamiento en los nodos afectan directamente al retardo en la red. [11]

A. Aplicación en el nodo Gateway WMG

El Gateway WMG, se configura de forma que la red implementada pueda soportar la pila de protocolos IPv6 y también para que se implemente el túnel que permita la comunicación hacia la red IPv6. La red 6LoWPAN usa direcciones públicas con 48 bits para la sección de los host. El Gateway WMG usa la librería mrv6-edge que posibilita al nodo cumplir las funciones de Gateway y por tanto armar la red y conectar la red IPv6. La configuración de este nodo, se la realiza en el Shell de Mote Runner que se ejecuta en la computadora que implementa el túnel IPv4/IPv6 con los comandos que se indican en la figura N° 7.



```
mote-create -i 192.168.1.223
02-00-00-00-8E-D7-28-1D
cd /moterunner/examples/mrv6/lib/js
/moterunner/examples/mrv6/lib/js
source mrv6.js
null
```

Fig. N° 7. Configuración del nodo WMG

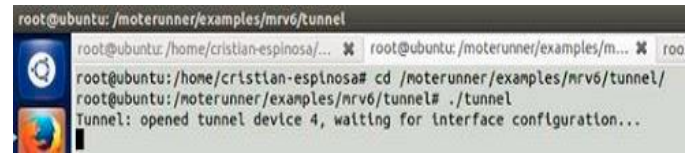
Para crear la dirección IPv6 de los nodos sensores, se utiliza la dirección MAC de las mismas. Para identificarlos se puede utilizar las direcciones de 16 bits proporcionadas por el Gateway WMG y direcciones de 64 bits que utiliza la dirección MAC del mote, mapeadas con direcciones IPv6 en el WMG. El computador que implementa el túnel, permite la transmisión de los datos a la red IPv6 y nos permite gestionar los nodos de la red 6LoWPAN, visualizando información propia de cada uno de ellos, tal como se muestra en la figura N° 8.



```
starting mrv6 wsn is not yet active.
network-list
No simulations started.
-----
Mote-Abbrev Name State Connection Uniqueid Address
-----
u6-w0 192.168.1.22 on IDP:192.168.1.223:9999 02-00-00-00-8E-D7-28-1D udp://192.168.1.223:9999
w1 Mote 0 on MRV6 02-00-00-00-93-41-08-13 ff://02-00-00-00-93-41-08-13 v6://02-00-00-00-93-41-08-13
w2 Mote 1 on MRV6 02-00-00-00-65-1F-EE-15 ff://02-00-00-00-65-1F-EE-15 v6://02-00-00-00-65-1F-EE-15
-----
v6-connect
Mote Parent Address Hops
-----
02-00-00-00-8E-D7-28-1D Gateway 0 0
02-00-00-93-41-08-13 02-00-00-00-8E-D7-28-1D 1 1
02-00-00-65-1F-EE-15 02-00-00-00-8E-D7-28-1D 2 1
```

Fig. N°8. Estado de los WM.

Este computador además se configura para realizar el enrutamiento de los datos que entran y salen de la red 6LoWPAN como se muestra en la figura N° 9. sobre la configuración del túnel.



```
root@ubuntu: /moterunner/examples/mrv6/tunnel
root@ubuntu: /home/cristian-espinoza/...
root@ubuntu: /moterunner/examples/m...
root@ubuntu: /home/cristian-espinoza# cd /moterunner/examples/mrv6/tunnel/
root@ubuntu: /moterunner/examples/mrv6/tunnel# ./tunnel
Tunnel: opened tunnel device 4, waiting for interface configuration...
```

Fig. N° 9. Configuración del túnel IPv4/IPv6

B. Aplicación en los nodos sensores WM

En cuanto a los nodos sensores WM, se debe cargar las aplicaciones necesarias para que puedan trabajar en IPv6 y manejen los sensores conectados a los mismos. La lectura del dato sensado y la transmisión hacia el WMG también deben ser configuradas. Las librerías de C# que se pueden usar para la programación de los WM son limitadas debido a la capacidad de procesamiento del nodo. Los tipos de datos que se pueden usar también son limitados. Para transmitir los datos desde nodos WM, se creó paquetes IPv6 con la respectiva cabecera y datos. Una vez creado el paquete, este se envía a través de un socket que se encuentra creado en el PC conectado en el otro lado de la red IPv6. Un aspecto crítico a considerar en la implementación de este tipo de redes es la velocidad de procesamiento de los nodos, pues esto influye en la capacidad de respuesta de la red como tal y debe tomarse en cuenta a la hora de escribir las aplicaciones.

C. Aplicación de procesamiento de datos

En el servidor se implementa el software que permite obtener la información de los sensores de los nodos por medio de un socket enlazado al socket de los WM. Parte del código se presenta a continuación:

```
Byte[] datosRx = servidor.Receive(ref equipoRemoto);
string direccionMo = "2005::200:0:0:20";
if
(equipoRemoto.Address.Equals(IPAddress.Parse(direccionCO
)))
{
int datos1 = BitConverter.ToInt32(datosRx, 0);
datos2 = bytesToLong(datosRx);
string C0strings = transV01(datos2);
string ppm = fPpm(datos2);
```

V. CONCLUSIONES

Luego de implementar el prototipo de Gestión de Luminarias LED, se puede concluir que la utilización de 6LoWPAN en redes WSN, es una buena opción frente a otras tecnologías alámbricas e inalámbricas que se han venido utilizando hasta el momento en las mismas.

Las limitaciones propias en el procesamiento de los nodos influyen drásticamente en el performance de IoT en 6LoWPAN y es el parámetro más influyente en el desarrollo de las redes que implementas soluciones usando WSN.

Cuando las comunicaciones son manejadas a nivel de capa de enlace se eliminan retrasos en la entrega de los datos, sin embargo esto implica que se tenga una sola interconexión con la red IPv6 y por tanto es un punto de fallo del sistema.

Si los nodos deben transmitir un alto volumen de datos el retardo en la transmisión debido al procesamiento de los mismos se convierte en un parámetro crítico para la solución implementada. En el caso del prototipo propuesto esta consideración no es crítica, puesto que estamos sensando si la luminaria está encendida o no.

REFERENCIAS

- [1] . Ajith Kumar S., Knut Ovsthus, and L. M. Kristensen "An Industrial Perspective on Wireless Sensor Networks.A Survey of Requirements, Protocols, and Challenges" IEEE Communications Surveys & Tutorials, Vol. 16, No. 3, 2014
- [2] Zigbee Document 053474r13, Zigbee Alliance, 2008
- [3] G. Mulligan. "The 6LoWPAN architecture," Proc. Embedded networked sensors, New York, 2007, pp. 78–82
- [4] A. Ajith Kumar S., Knut Ovsthus, and L. M. Kristensen "An Industrial Perspective on Wireless Sensor Networks.A Survey of Requirements, Protocols, and Challenges" IEEE Communications Surveys & Tutorials, Vol. 16, No. 3, 2014
- [5] <http://www.libelium.com>
- [6] "Moterunner"<http://www.zurich.ibm.com/moterunner/>
- [7] G. Montenegro, N.Kushalnagar, J Hui and D culler, "Transmission of IPv6 Packets over IEEE 802.15.4 Networks" RFC 4944, September 2007
- [8] "Wasmote runner"<http://www.libelium.com/products/wasmote-mote-runner-6lowpan>
- [9] "Wasmote Mote Runner Technical Guide" Document version: v4.0 - 10/2013
- [10] "Informe de Gestión", Empresa Eléctrica Quito Diciembre, 2016
- [11] Fadi M. Al-Turjman Ashraf E. Al-Fagih Waleed M. Alsalih Hossam S. Hassanein. "A delay-tolerant framework for integrated RSNs in IoT". Computer Communications, Volume 36, Issue 9, 15 May 2013, Pages 998–1010.

Implementación de un aplicativo para verificar la conectividad de nodos de sensores inalámbricos utilizando IPV6

Luis M Yulán¹ Egas A. Carlos¹ Aguirre V. Vanessa¹ Casierra, Juan²

luis_enriqueym@hotmail.com, carlos.egas@epn.edu.ec, vanessa.aguirre@epn.edu.ec, casierrajuan@gmail.com

¹Escuela Politécnica Nacional- Ecuador

¹Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador

Resumen— Las redes sensores inalámbricas tienen limitaciones en el consumo de energía y capacidad de procesamiento, esto ha causado que muchas investigaciones tengan como objetivo minimizar el consumo de energía en los nodos sensores. Gran parte del consumo de energía se da cuando los nodos están en modo de transmisión y recepción, adicionalmente, el consumo de energía cuando los nodos están procesando información puede resultar importante. La gestión remota de la red sensor inalámbrica, involucra procesamiento adicional en los nodos y transmisión de tramas de gestión lo cual afecta al consumo de energía de los nodos sensores. Una de las principales tareas en la gestión de redes sensores inalámbricas, es la verificación de la conectividad en la red, lo cual se ve obstaculizado cuando los nodos sensores están en modo dormido, modo en el cual no pueden transmitir y recibir tramas de gestión y por lo tanto responder a mensajes ICMPv6. En este artículo nosotros proponemos un método para verificar la conectividad de nodos sensores utilizando IPV6 minimizando el consumo de energía de los nodos sensores, dejando gran parte de los procesos de gestión al Gateway y resolviendo el problema de verificar la conectividad de un nodo cuando está en el modo dormido.

Palabras clave— Redes sensores inalámbricas, sistema operativo TinyOS, prueba de conectividad, protocolo IPV6

I. INTRODUCCIÓN

Las redes WSN (Wireless Sensor Network) en los últimos años han tenido un rápido crecimiento a tal punto que las podemos encontrar en diferentes campos de la ingeniería, debido a sus múltiples aplicaciones (domótica, monitorización ambiental, cuidados en la salud, etc.). Esto ha convertido a las WSN en una tecnología ideal para el desarrollo del Internet de las cosas (Internet of Things, IoT).

Los nodos en una WSN están típicamente organizados en tres tipos de topologías como se indica en la figura N° 1. Topología de estrella, donde cada nodo se conecta directamente al nodo Gateway (Puerta de enlace). Topología de árbol, donde cada nodo se conecta a un nodo de mayor jerarquía en el árbol y después al nodo Gateway. Finalmente, para ofrecer mayor confiabilidad existe la topología tipo malla en donde los nodos se pueden conectar a múltiples nodos en el sistema y pasar datos por el camino disponible de mayor confiabilidad [1]. En todas ellas existe al menos un nodo coordinador que conoce la existencia de todos los nodos sensores en la red dicho nodo por lo general es el nodo Gateway.

Una actividad básica en la gestión de la red, es la verificación de conectividad de los nodos, situación que se ve reflejada con la implementación del comando ping en todos los sistemas operativos. Por otra parte las implementaciones de las redes sensores inalámbricas, han influido en el desarrollo de protocolos a medida, para solucionar los requerimientos

específicos, con el propósito de minimizar los problemas del retardo, limitaciones de energía y poca capacidad de procesamiento de los nodos sensores. Se sigue trabajando en la IETF en un protocolo que permita realizar la gestión de red en las diferentes implementaciones de WSN [2].

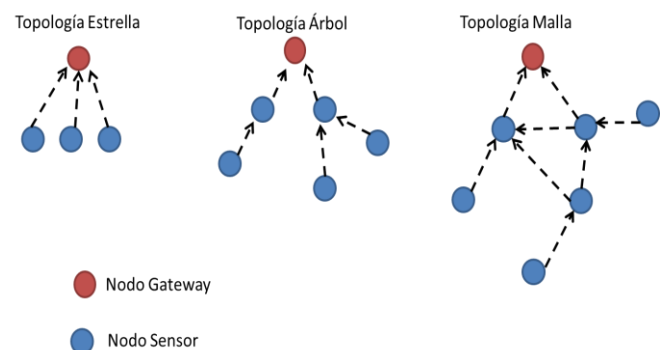


Fig. N°1 Topologías de red WSN

Dentro de las redes WSN se distinguen tres elementos principales:

- **Nodo Sensor:** Son los encargados de recolectar la información deseada y procesarla para transmitirla por medio de la red hacia el destinatario.
- **Nodo Gateway:** Elemento encargado de interconectar la red de sensores con otros dispositivos que se encargarán de procesar e interpretar los datos provenientes de la WSN.
- **Estación base:** Agente encargado de recolectar los datos que llegan del Gateway permitiendo al usuario monitorizar la red tanto de manera local o remota

Como elementos secundarios podemos tener a nodos coordinadores o interceptores que se encargan de replicar la información de un nodo sensor hasta el nodo Gateway. En la actualidad existen implementaciones IPv6 sobre el estándar IEEE 802.15.4. Por ejemplo, 6LoWPAN [3], que utiliza un direccionamiento con menor número de bits con el fin de optimizar el tamaño máximo de la PDU de IEEE 802.15.4 que es igual a 128 bytes.

En una red WSN que utiliza 6LoWPAN, para verificar la conectividad de los nodos sensores de forma remota, se requiere que los nodos sensores tengan una dirección IPv6, y se lleva a cabo principalmente el protocolo ICMPv6 para lo cual el nodo necesita estar en estado activo y de esta manera pueda ser administrado, sin embargo en las redes WSN los nodos que forman parte de la red WSN, pueden estar en modo dormido, de bajo consumo, donde no están activados los procesos de recepción y transmisión, por lo tanto, son incapaces de recibir y procesar peticiones ICMPv6, generando la conclusión errónea de que el nodo no está

operativo. Esta condición es un problema a resolver en la gestión de WSN, el cual lo resolvemos con nuestra propuesta.

El consumo de energía en los nodos puede ser minimizado, si el encargado de responder las peticiones ICMPv6 fuera la estación base, la cual no tiene limitaciones en la disponibilidad de energía. Por lo tanto, la estación base optimizaría el consumo de energía de un nodo sensor al ser el encargado de responder al cliente remoto.

Por otra parte, la estación base es la única que conoce el verdadero estado de los nodos en la WSN que puede ser: modo de transmisión, recepción o ahorro de energía, en base a ese conocimiento la estación base evita retransmisiones innecesarias hacia un nodo determinado, esperando el momento adecuado para hacerle llegar la información correspondiente.

II. TRABAJOS RELACIONADOS

La forma más común que se utiliza para verificar la conectividad de los nodos es utilizando el comando ping, herramienta implementada en la mayoría de los sistemas operativos para WSN tales como TinyOS [4], Contiki [5], los cuales tienen en cuenta las limitaciones de memoria, capacidad de procesamiento y la disponibilidad de los nodos sensores.

Para la implementación del comando ping con IPv6, en el sistema operativo TinyOS [6], el cual es un sistema operativo de código abierto desarrollado para trabajar con sistemas embebidos que permite cumplir con las restricciones de memoria que necesitan las redes de sensores inalámbricos, es necesario la utilización del módulo de BLIP [7], que adapta el protocolo IPv6, para el uso en WSN. Esto implica líneas adicionales que se ejecutarán consumiendo de los recursos de energía, memoria y procesamiento en el nodo que implemente dicho código. Lo mismo sucede con otros sistemas operativos como Contiki, que utiliza IPv6 [8], eCos [9], Riot [10], EYESOS [11], Mantis [12] y Duinos [13] los cuales requieren verificar la conectividad utilizando el comando ping.

En el caso de los sistemas operativos tales como MagnetOS [14], Nano -RK [15] y Cromos [16] que no han implementado 6LoWPAN y por lo tanto no es posible verificar la conectividad remota a través de una red IPv6.

Nuestra propuesta consiste en dar la responsabilidad de responder las peticiones ICMPv6 realizadas a cualquier nodo en la red WSN, a la estación base de la red WSN, para que responda a dichas peticiones, dejando al nodo sensor la única labor de procesar datos y estando mucho más tiempo en el modo ahorro de energía. La estación base puede hacer esto gracias al trabajo que realiza en conjunto con el nodo Gateway que actúa como coordinador de la red, y que tiene toda la información de los nodos de la red, información necesaria para conocer el estado de un nodo sensor. Para que la estación base pueda realizar la contestación ICMPv6 como si fuese un nodo, se realizó la asignación de interfaces virtuales IPv6 a cada nodo en la estación base. La interfaz virtual es dada a un equipo remoto que monitoreará el estado de los nodos en base a dicha interfaz virtualizada

III. IMPLEMENTACIÓN

Ponemos en práctica la solución propuesta, implementando una red WSN usando TinyOS [4]. Para agregar un nodo o identificar si este está activo, capturamos los datos que

circulan en la red y por lo tanto extraemos la información necesaria de la trama de datos (payload) que genera cada nodo, específicamente su identificador, y de esta manera verificar la conectividad de los nodos, sin necesidad de utilizar el comando ping existente en la red 6LoWPAN. Así nuestra propuesta, con modificaciones menores se puede utilizar para verificar la conectividad de los nodos sensores en diferentes sistemas operativos.

TinyOS está diseñado para funcionar con el estándar IEEE 802.15.4, pero definiendo sus propias sub-capas, por ejemplo a la capa MAC es sustituida por una llamada capa B-MAC [17] como se indica en la figura N°2, debido a que TinyOS utiliza abstracción de hardware, por lo que es una manera eficiente e ideal para sistemas con redes de sensores.

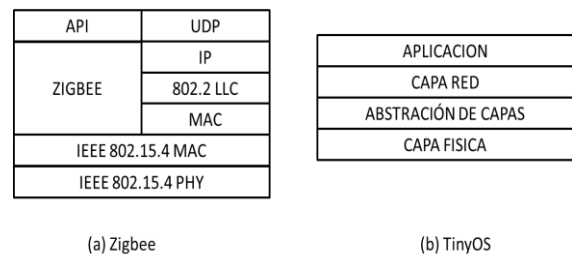


Fig. N° 2 Modelo de capas (a) ZigBee (b) TinyOS.

Para la elaboración de nuestra red multisalto en TinyOS, hemos seleccionado el protocolo CTP (Collection Tree Protocol) porque, entre sus características no es dependiente de la plataforma (tanto microcontrolador y radio). También tiene un rendimiento aceptable con respecto a otros sistemas multihop desarrollados para nodos sensores inalámbricos.

Este algoritmo selecciona al nodo Gateway como el nodo raíz o sink (colector). CTP usa parámetros denominados ETX. En donde el nodo raíz que vendría a ser el Gateway tendría un ETX=0; un nodo cualquiera tendrá un valor de ETX inmediatamente por encima de él (llamado nodo padre) más el valor ETX del enlace entre ellos. Es decir, cada nodo decidirá la mejor ruta de acuerdo al ETXmin [18].

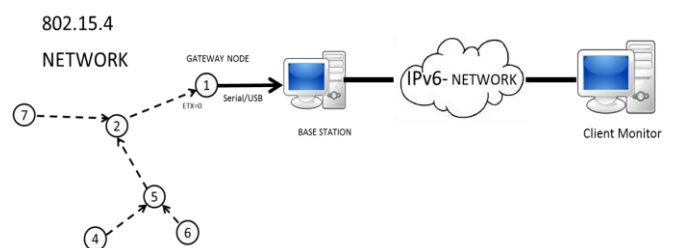


Fig. N° 3 Prototipo Implementado

El nodo Gateway de la figura 3 conectado a la estación base a través del puerto de serie, actúa como coordinador de la red WSN, por lo tanto, se encarga de asociar los nodos a la red y es responsable de la captura de las tramas para de esta manera, obtener información sobre los nodos sensores que conforman la red y conocer cuales se encuentran activos. Esta información se envía a la estación base que ejecuta la aplicación colectora, que se encarga de responder a las peticiones del cliente remoto (aplicación monitora) que verificará el estado de los nodos de sensores que utilizan una red IPv6. El cliente conectado a la red IPv6 que requiere verificar la conectividad de los nodos de manera remota, realiza un ping a la interfaz virtual asociada a cada nodo, que se encontrará activa si el nodo se encuentra activo.

La red WSN utiliza nodos sensores IRIS XM2110CB [19], que están programados en NesC lenguaje de programación de TinyOS. Para desarrollar del software que se ejecutó en la estación base y cliente remoto, Se utilizó JAVA-Eclipse que permite crear las interfaces gráficas para un control más simple de la red.

Las aplicaciones diseñadas en cada una de las etapas del sistema se enlistan a continuación:

- **Aplicación Envío.C:** se encuentra dentro de la primera etapa (red WSN) y es utilizada por los nodos para enviar datos censados hacia su nodo Gateway.
- **Aplicación Base_root:** es diseñada para el nodo Gateway y se encarga de recibir y enviar datos de la red WSN a la aplicación monitora mediante un puerto serial.
- **Aplicación Colectora:** esta aplicación realiza la función mencionada en la segunda etapa, pero es cargada en una PC con sistema operativo Linux.
- **Aplicación Monitora (Ping):** se encarga de mapear las direcciones IPv6 de cada nodo sensor de forma remota cumpliendo con lo mencionado en la etapa cuatro. Esta aplicación puede utilizarse en sistemas operativos Windows y Linux.

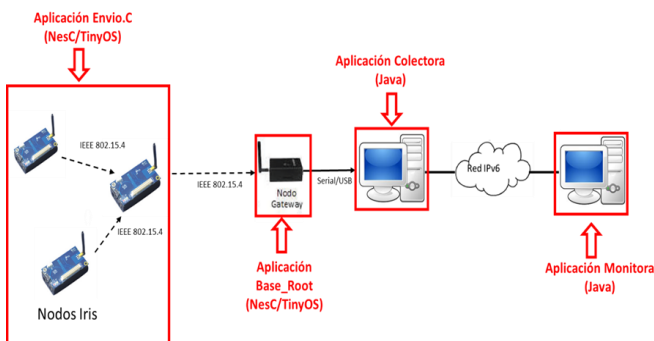


Fig. N° 4 Esquema del sistema implementado

A. Aplicación Envío

Esta aplicación permite tener enrutamiento de paquetes entre nodos, consumo eficiente de energía y la realización de lecturas de voltaje y temperatura. El algoritmo que ejecuta esta aplicación es mostrada a continuación:

Algoritmo 1 Envío de datos a la red WSN

- 1 Salir del modo LPL (Low Power Listening)
- 2 Lectura de temperatura
- 3 Medición del nivel de voltaje
- 4 **Timer** (envió datos) Extrae los datos leídos para enviarlos a la red WSN
- 5 De regreso al modo LPL

B. Aplicación Base_Root

El nodo Gateway responsable de recoger toda la información de los nodos sensores que estén activos en la red WSN, esta información recolectada se envía por la conexión serie a la estación base que suele ser un ordenador o PC que contiene la aplicación Colectora. Este nodo siempre está en modo escucha con el fin de informar a la estación base sobre la actividad de la red y también listo para asociar nuevos nodos a la red WSN. A continuación se muestra el algoritmo usado en el nodo Gateway.

Algoritmo 2 Recepción del Nodo Gateway

- 1 Inicialización de protocolo CTP.
- 2 Posicionamiento del nodo Gateway como nodo Raíz (setRoot)
- 3 Asociación de nodos sensores.
- 4 **While** (Escucha Transmisiones)
- 5 Escucha de transmisiones para recepción de datos por parte de los nodos activos o nuevos nodos
- 6 Recibe información por la interface radio
- 7 Transmite información del nodo activo a la estación base vía interfaz serial
- 8 **End while**

Para asignar un identificador a cada nodo sensor, se debe conectar mediante puerto serial la placa MIB520CB acoplado al nodo que se desee cargar la aplicación y ejecutar el segmento de código dentro de la carpeta de la aplicación a cargar en una terminal de Linux:

```
make iris install, N mib520, /dev/ttyUSBx
```

La variable N, representa el identificador que se da al nodo sensor, mientras que MIB520 es la placa programadora que permite transmitir el código por el puerto serial x, que se conecta a dicha placa. El identificador asignado al nodo Gateway de esta red fue 1; ya que es el ID por defecto.

C. Aplicación Colectora

El concepto de crear o configurar varias direcciones IP en una única interfaz de red se denomina Asignación de alias IP y es muy útil para configurar varios sitios virtuales en Apache utilizando una única interfaz de red con direcciones IP diferentes. La ventaja principal del uso de esta técnica permite que una sola interfaz física tenga varias IP, utilizando muchas interfaces virtuales [21]. Esta técnica permite la asignación de direcciones IPv6 en base a la ID de cada nodo sensor.

La estación base está conectada a la red WSN y la red IPv6 como se indica en la figura N° 4, asegurando de que los nodos sensores sean visibles a redes externas o en este caso a la red de área local de la aplicación monitora. Esta aplicación permite la asignación de interfaces virtuales con direcciones IPv6 a cada uno de los nodos sensores. El estado de la interfaz virtual creada es el mismo que el estado del nodo sensor (ON/OFF). La estación base asocia el ID del nodo con una dirección IPv6 para evitar conflictos de duplicidad. Esta información es compartida a la aplicación monitora denominada *ping*, para que esta pueda conocer el estado de los nodos sensores.

Las tramas que llegaran a la interfaz serial deben ser interpretadas de la forma correcta, para ello es necesario saber los tipos de tramas que utiliza TinyOS como se indica en la figura N°5. Los datos que nos interesan en la trama de datos de TinyOS son el ID del sensor y los 6 bytes que tendrá el payload del mensaje.

La aplicación colectora en la estación base comprueba si el nodo pertenece a la red y evita que los nodos fuera de la red se incluyan, proporcionando el mecanismo para eliminar un nodo sensor de la lista.

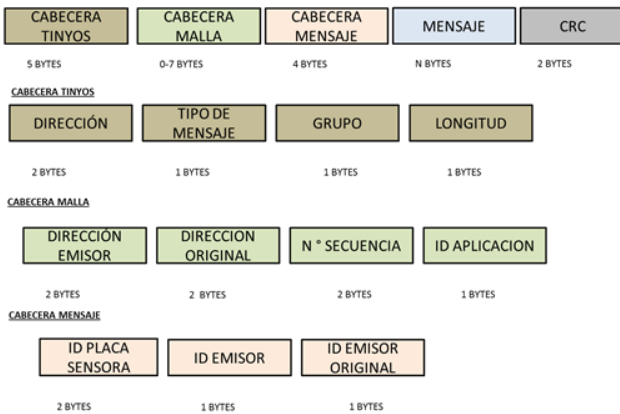


Fig. N° 5 Tipos de Trama TinyOS para una red multihop

Para extraer el ID de cada nodo se utilizó la librería RXTX de java [22]. Para la lectura de los datos del puerto serial, se debe sincronizar con el nodo Gateway a una velocidad de sincronización igual a 57600 simb/seg. Una vez asignadas las direcciones y el ID asociado con la dirección IPv6, la información está lista para ser enviada a la estación de cliente o para responder a un ping.

A continuación se muestran los algoritmos usados por la aplicación colectora que se ejecuta en la estación base:

Algoritmo 3 Obtención de información del estado de los nodos sensores

- 1 Creación de tabla para almacenamiento de ID, dirección IPv6, datos de voltaje y temperatura
- 2 **mitimer** timer para setear datos a off.
- 3 **tumbral** tiempo máximo de espera para datos
- 4 **valor_serial** variable para captura de valores
- 5 Inicializamos el timer con el tiempo de umbral **mitimer.start(tumbral)**
- 6 **WHILE**(escucha puerto serial)
- 7 **If** (valor_serial >0)
- 8 Reinicializamos el Timer con el tiempo umbral
- 9 Capturamos los datos, su ID y los almacenamos
- 10 Presentamos los datos en Tabla
- 11 **Endif**
- 12 **End while**

Algoritmo 4 Comunicación por socket con la aplicación monitora ping

- 1 Extracción de datos
- 2 **Envio_datos**(envía datos por sockets)
- 3 Abertura del puerto socket
- 4 Serialización de datos a enviar
- 5 Enviamos datos de tabla por medio de sockets a la IP del cliente
- 6 Cierre del puerto socket
- 7 **End Envio_datos**



Fig. N° 6 Estación base en funcionamiento con la aplicación monitora Ping

D. Aplicación Monitora Ping

En el caso del cliente monitor ping, se ha desarrollado una aplicación implementada en la estación base indicada en la figura N°6, que permite obtener la tabla de datos de los nodos de sensores activos y no activos que forman parte de la red WSN, y hacer ping directamente la dirección IPv6 virtual de un nodo sensor específico.

Algoritmo 5 Cliente generación de Ping

- 1 **tenvio** para pedir información cada cierto tiempo
- 2 **timercliente** para enviar paquete al socket de la estación base
- 3 **WHILE** (comunicación por sockets)
- 5 Inicialización del timer con el tiempo de envío **timercliente.start(tenvio)**
- 6 **if** (chequeo de bandera del timer si llega al límite)
- 7 Envía paquete para pedir conexión
- 8 Receptada datos del servidor socket que será la estación base
- 9 Armamos la tabla y adicionalmente se tendrá un botón que permita hacer ping a una dirección IPv6 de la tabla asociada a los nodos
- 10 **End while**

IV. PRUEBAS REALIZADAS

Se realizaron pruebas individuales para cada una de las cuatro etapas del sistema:

1. Prueba de funcionamiento de la red WSN con TinyOS.
2. Recepción, direccionamiento y envío de datos por parte de la aplicación colectora.
3. Enrutamiento de la red IPv6 creada.
4. Funcionamiento de la aplicación monitora ping.

La prueba total del sistema se realizó en la aplicación monitora ping, que abarca toda la funcionalidad del sistema. Luego de comprobar que la aplicación ping recibía datos, se procedió a comprobar que el botón de ping funcione de acuerdo a lo requerido. En la figura N° 7, se puede visualizar el estado de la conectividad de los nodos sensores y las respuesta a la ejecución del comando ping para la dirección IPv6 seleccionada y que concuerda con el estado en que se encuentra el nodo.

Como resultado del presente trabajo de investigación, se implemento la aplicación que permite verificar la conectividad de nodos sensores en una red 802.15.4 con IPv6, minimizando los procesos a realizar en el nodo sensor y por lo tanto minimizando el consumo de energía, permitiendo establecer la conectividad de los nodos sensores,

aun cuando están en el modo dormido, y se pudo comprobar el buen funcionamiento de gestión centralizada de la red, dejando la mayor parte de responsabilidades al nodo Gateway.

V. CONCLUSIONES

La herramienta propuesta en este documento puede ser utilizado eficientemente en redes 6LoWPAN para verificar los nodos sensores de conectividad. El sistema desarrollado, mejora la capacidad de detectar la conectividad del nodo sensor cuando el nodo está en estado de reposo, lo que minimiza las respuestas de nodo caído.

ID de Nodo	Direccion IPv6	Estado	Voltaje [V]	Temperatura [C]
16	2001:0:0:1::16	estado ON	2.6769156	12.6384
17	2001:0:0:1::17	estado ON	2.6769156	12.6636
19	2001:0:0:1::19	estado ON	2.7979822	12.546

Fig. N°7 Estado de conectividad de los nodos en la aplicación monitora ping

En el caso de que ninguna red WSN ha implementado 6LoWPAN, la herramienta proporciona un método para verificar los nodos de conectividad utilizando la red IPv6.

En WSN con el protocolo 6LoWPAN, la herramienta desarrollada impide ejecutar proceso adicional en el nodo sensor para verificar la conectividad.

La realización de este trabajo, aumenta el debate relacionado con la decisión de tener un sistema de gestión de los nodos de sensores distribuidos o centralizados en el nodo de pasarela, teniendo en cuenta las limitaciones de la energía en los nodos

Con la virtualización de direcciones IPS, se puede tratar a cada nodo como un elemento independiente, creando varios aplicativos para ser tratados de forma distinta y a su vez que cada uno transmita por socket varios datos gracias a la dirección IP dada.

Finalmente, consideramos que este trabajo permite la verificación de la conectividad de los nodos en una ambiente IPv6 y es el inicio del desarrollo de aplicaciones de gestión de redes para 6LowPAN. Queda pendiente la realización de la medición del consumo de energía de la aplicación desarrollada y su comparación con otras técnica utilizadas.

VI. RECOMENDACIONES

A pesar de que no se utilizó en su totalidad la tarjeta integrada MTS400 en los nodos iris, se puede hacer uso de las demás funciones, tales como: medición de humedad, ruido y movimiento; dando apertura a un rango más grande de usos para la red WSN creada. Para ello se aumentará en la programación de EnvioC.nc, el llamado a dichas interfaces

con call; cabe recalcar que mientras más funciones e interfaces se utilice, más rápido se consumirá la energía de las baterías.

Es necesario crear con la utilización de sockets un ambiente multiplataforma, ya que la recepción de los datos se puede dar en cualquier tipo de interfaz. Por ejemplo, se podría programar en Android Studio para que el monitoreo de datos sea a través de un dispositivo celular siempre y cuando exista conectividad en dicho dispositivo.

REFERENCIAS

- [1] National Instrument, «Que es una red de sensores inalámbricos,» National Instrument, 15 Abril 2009. [En línea]. Available: <http://www.ni.com/white-paper/7142/es/>.
- [2] A. H. Jose Araujo, «Inverted Pendulum Control over an IEEE 802.15. 4 Wireless Sensor and Actuator Network,» 23 Febrero 2011. [En línea]. Available: https://www.researchgate.net/publication/264005006_Inverted_Pendulum_Control_over_an_IEEE_802154_Wireless_Sensor_and_Actuator_Network.
- [3] C. R. C. Peña, «IPv6: The engine of "The Web of Things",» blogthinkbig, 20 Abril 2013. [En línea]. Available: <http://blogthinkbig.com/ipv6-motor-internet-de-las-cosas-iot/>.
- [4] P. Levis, Tinyos-Programing, California: Barkeley, 2006.
- [5] R. Fuentes, «Contiki O.S para IoT,» rfuentess.blogspot, 04 Abril 2016. [En línea]. Available:<http://rfuentess.blogspot.com/2016/04/contiki-os-para-el-iot.html>.
- [6] V. Asanza, «Tinyos en redes de Sensores,» Academia, 16 Junio 2010. [En línea]. Available: https://www.academia.edu/15075955/TinyOS_en_redes_de_sensores. [Último acceso: Agosto 22 2017].
- [7] Tinyos Community, «BLIP_TUTORIAL,» TinyOSWEB, 15 Octubre 2011. [En línea]. Available:http://tinyos.stanford.edu/tinyoswiki/index.php/BLIP_Tutorial.
- [8] Contiki O.S Community , «Pv6 specific features,» Contiki O.S Community , 20 Julio 2012. Available: <http://contiki.sourceforge.net/docs/2.6/a01790.html>.
- [9] Ecos Home Pages, «Contributions and developmente projects,» Ecos Home Pages, 07 Noviembre 2015. [En línea]. Available: <http://ecos.sourceforge.org/>.
- [10] Riot Home pages, «Features about Riot,» Riot Home , 06 Junio 2016. [En línea]. Available: <http://riot-os.org/>
- [11] EYESOS, «Software-general,» recursostic, 08 Diciembre 2012. [En línea]. Available: <http://recursostic.educacion.es/observatorio/web/fr/software/software-general/942-eyesos>.
- [12] Project MANTIS, «Sensor system MANTIS,» Project MANTIS, 15 Abril 2012. [En línea]. Available:<http://www.cs.colorado.edu/~rhan/sensornets.html>.
- [13] Arduino, «Arduino IPv6 stacks,» TELECOM , 15 Noviembre 2015. [En línea].Available: <https://departements.telecom-bretagne.eu/rsm/logiciels>.
- [14] MAGNETOS Community, «MAGNETOS Related Work,» MAGNETOS Community, 06 Septiembre 2015. [En línea]. Available: <https://www.cs.cornell.edu/people/egs/magnetos/related.html>.
- [15] Nano- RK, «Nano- RK Wiki,» Nano- RK, 04 Abril 2011. [En línea]. Available: <http://nanork.org/projects/nanork/wiki>.
- [16] CORMOS , «A Communication Oriented Runtime System for Sensor Networks,» CORMOS Seminar, 20 Julio 2005. [En línea]. Available: <http://osnet.cs.nchu.edu.tw/powpoint/seminar/2005/2005%20CORMOS.pdf>.
- [17] J. Polastre, «Sensor Network Media Access Design,» Univ. de California Berkeley, California, 2003.
- [18] J. Tarifa, "Red de sensores inalambricos para la monitorización de instalaciones eléctricas de baja tensión" tesis de pregrado, Sevilla: Universidad de Sevilla, 2003.

- [19] MENSIC INC, MTS/MDA Sensor Board Users Manual, San Jose: Crossbow Technology, 2012.
- [20] Grupo de investigacion EDMANS, Redes sensores inalambricas: teoría y aplicación práctica, Buenos Aires: Univ. de la Rioja, 2005.
- [21] N. Shrestha, "Create Multiple IP Addresses to One Single Network Interface" Tecmint, 02 Julio 2013. [En línea]. Available:<http://www.tecmint.com/create-multiple-ip-addresses-to-one-single-network-interface/>
- [22] http://www.jcontrol.org/download/rtx_en.html

Simulador de cognitive radio en redes móviles

Palacios, Pablo¹, Jonathan, Delgado²
 pablo.palacios@ug.uchile.cl, jonadelceing84@gmail.com,

¹Universidad de Chile Santiago-Chile

²Universidad Técnica de Manabí Manabí-Ecuador

Resumen— En este trabajo se prueba e implementa métodos y algoritmos matemáticos aplicados a la tecnología de Cognitive Radio (CR), mediante el desarrollo de un simulador CR aplicado a las redes inalámbricas móviles WiFi y LTE, estructurado con cada una de las cuatro etapas de un sistema CR, usando herramientas matemáticas evaluadas en trabajos previos de manera individual. La implementación del simulador CR adapta de manera conjunta a distintos algoritmos, los cuales son: Singular Value Decomposition (SVD), Teoría de juego de coaliciones y Handoff de tecnologías por RSSI. Se usa el software Network Simulator 3 (NS-3) como entorno de desarrollo y ejecución de una red cognitiva heterogénea basada en el simulador creado, permitiendo evaluar el rendimiento de los métodos aplicados a la red móvil CR mediante resultados numéricos comparables con otros algoritmos aplicables a CR, logrando validar un análisis teórico previo en una red móvil funcional.

Palabras clave: Cognitive Radio, Redes Móviles, Simulador CR, Singular Value Decomposition (SVD), Teoría de Juegos, RSSI, NS-3.

I. INTRODUCCIÓN

El análisis del desempeño de Cognitive Radio (CR) genera un gran interés en las comunidades académicas e industriales para satisfacer las crecientes necesidades de los recursos de espectro y la comunicación de datos de alta velocidad.

Aguilar y Navarro [1] dan una reseña sobre la tecnología conocida como Cognitive Radio (CR), que ha evolucionado con el transcurso del tiempo, desde llamarse en una primera instancia Software Radio (SR), la cual era una radio reconfigurable, capaz de realizar distintas funciones en tiempos diferentes, únicamente realizándole cambios en su configuración mediante software. Luego paso ser Software-Defined Radio (SDR), que no era otra cosa que una versión de SR implementable y mejor estructurado. Una de las primeras pruebas de SDR fue el proyecto militar estadounidense SpeakEasy que implementó diez tipos de tecnologías inalámbricas en un rango de 2 Mhz hasta los 200 Mhz, en un solo equipo reprogramable.

Todo esto se llevó a cabo hasta llegar a 1999, Donde Joseph Mitola III [2] definió a CR como un sistema SDR inteligente, capaz de detectar necesidades de comunicación de los usuarios, poder conocer su entorno y gracias a este conocimiento llegar a satisfacer las necesidades de dichos usuarios.

En la actualidad no existen redes CR desplegadas comercialmente, a pesar de que las ideas del sistema de compartir diferentes tecnologías sin producir interferencia ya ha sido demostrada. Se han realizado pruebas en la banda de los 5 GHz de 802.11a para ser compartida con los radares de gran sensibilidad de navegación aeronáutica, también se han realizado pruebas e incluso estandarizado la tecnología para

tener acceso dinámico al espectro de TV en la banda entre 400-800 Mhz lo que se conoce como Cognitive Radio Wireless Regional Area Network (WRAN) y su estándar IEEE-802.22, como especifica Kalil [3].

En los últimos diez años ha existido un significativo interés en la investigación de CR en cada una de sus etapas, tanto en la academia como en la industria, presentando muchas propuestas, métodos y algoritmos para su implementación en un sistema CR, pero hasta el momento una arquitectura con métodos estandarizados no ha podido ser definida, como indica Cabric [4].

La principal ventaja que ofrece CR es la eficiencia en el uso del espectro electromagnético, ya que permite una gestión óptima del espectro a través de un proceso de cuatro etapas (ciclo cognitivo): detección del espectro, decisión del espectro, compartición del espectro y movilidad. Dentro de todo el ciclo cognitivo, las etapas más críticas son las de detección y decisión, siendo estas áreas de enorme importancia para todo el proceso cognitivo y presentan dificultades al tener mayores exigencias de hardware para la detección adecuada de lo Primary Users (PU) y potenciales problemas de seguridad en la red. Bajo estos parámetros, la detección correcta y efectiva de los PU se convierte en una prioridad. Por lo tanto, se han propuesto varios métodos, incluyendo algoritmos de detección, esquemas de detección de energía, detección basada en el valor propio, detección por descomposición de valores singulares SVD, entre otros.

En este proyecto se desarrolla un software simulador CR aplicado a las tecnologías móviles WiFi y LTE, estructurado con cada una de las cuatro etapas de un sistema CR, para lo cual en cada una de estas etapas se usan métodos y algoritmos evaluados en trabajos previos de manera individual. La implementación del simulador CR adapta de manera conjunta a los distintos métodos, los cuales son: Singular Value Decomposition (SVD), Teoría del juego de coaliciones y handoff por RSSI. Se usa el software Network Simulator 3 (NS-3) como entorno de desarrollo de una red cognitiva heterogénea basado en el simulador creado, permitiendo evaluar el rendimiento de los métodos aplicados a una red móvil CR, esto permite validarlos y compararlos en una red móvil funcional.

En este artículo se explica la metodología utilizada, el diseño e implementación del simulador, los resultados generales obtenidos y finalmente se presentan las conclusiones.

II. DESARROLLO

1. Metodología

Para lograr los objetivos planteados en el proyecto, la metodología usada fue de investigación, desarrollo, simulación y análisis. Mediante la investigación de textos científicos (papers), que se explicarán en este apartado, se pudo escoger métodos basados en algoritmos matemáticos y de teoría de juegos para cada etapa del sistema CR. El flujo de la metodología investigada se muestra en la Figura N°1, y sigue al flujo de un sistema CR.

2. Etapa de detección: Detección por Singular Value Decomposition (SVD) para redes CR

Para la detección del espectro se debe usar el método de “SVD (Singular Value Decomposition) Detection” propuesto por Fahim y Raenn [5], con sus respectivas adaptaciones a este trabajo, el cual se describe a continuación:

- 1) Seleccionar el número de columnas de la matriz de covarianzas L en la interfaz receptora CR de los usuarios secundarios, tal que $k < L < N-k$, donde N es el número de muestras a tomar por parte del receptor y k es el número de valores singulares, para esta procedimiento considerar $k=2$ y $L=16$.
- 2) Proceder a obtener la matriz de covarianzas

$$R(N_s) = \frac{1}{N_s} \sum_{n=16}^{16-1+N_s} x(n)x^t(n) \quad (2.1)$$

Donde $x(n)$ es la SNR del canal que el receptor está censando en ese instante y dividida para el número de muestras.

- 3) Factorizar la matriz de covarianzas
- 4) Obtener los máximos y el mínimos eigenvalores de la matriz de covarianza $R(N_s)$ los cuales serán λ_{\max} y λ_{\min} .
- 5) Calcular el valor umbral para comparar con los eigenvalores con la siguiente fórmula:

$$\gamma = \frac{(\sqrt{N_s} + \sqrt{16})^2}{(\sqrt{N_s} - \sqrt{16})^2} * \left(1 + \frac{(\sqrt{N_s} + \sqrt{16})^{\frac{2}{3}}}{(N_s * 16)^{\frac{1}{6}} * (F_1^{-1}(1 - P_{fa}))} \right) \quad (2.2)$$

Donde P_{fa} es la probabilidad de falsa alarma que se requiere que sea ≤ 0.1 . La función $(F_1^{-1}(1 - P_{fa}))$ es la función Tracy-Widom que es distribución de probabilidad normalizada para los eigenvalores cuya distribución se muestra en la Tabla N° 1.

Tabla N°1. Función Tracy-Widom

t	F1(t)
-3.9	0.01
-3.18	0.05
-2.78	0.10
-1.91	0.30
-1.27	0.50
-0.59	0.70
0.45	0.90
0.98	0.95

- 6) Comparar la relación de eigenvalores máximos y mínimos de la matriz de covarianza con el umbral, por lo tanto Si $\frac{\lambda_{\max}}{\lambda_{\min}} > \gamma$, la señal está presente, de otro modo la señal no está presente.
- 7) Obtener el bit de detección de señal de usuario primario individual.

3. Etapa de decisión: Juego de Coaliciones para detección del espectro en CR

En esta etapa de decisión es donde se implementó la teoría de juegos basada en coaliciones (juego cooperativo) analizados por Saad, Han y Debah en [6] y por Wang, Son, Han y Saad en [7]. De todas los algoritmos y métodos analizados previamente, se decidió usar un tipo de detección de espectro colaborativo entre los SU debido a que gracias a estudios previos por Zhu, Wu y Zhou [8] se muestra una significativa mejora del rendimiento, aunque se debe tomar en cuenta que existe un tradeoff entre las ganancias en términos de probabilidad de detección del PU y los costes en términos de probabilidad de falsa alarma. Es decir aumentará la probabilidad de detección de un PU, pero también aumentará la probabilidad de falsa alarma. El algoritmo desarrollado se muestra a continuación:

- 1) **FASE 1:** Detección local, donde cada SU individual obtendrá su bit de detección de señal de PU.
- 2) **FASE 2:** Formación de coaliciones adaptativas, durante la formación de la coalición adaptativa se asume que cualquier usuario secundario de manera aleatoria puede comenzar el proceso de unión. Se forman las coaliciones basadas en el algoritmo de unión y separación indicado a continuación:
 - **Unión:** la coalición decide unirse siguiendo los siguientes pasos:
 - a) Se decide unir cualquier set de coaliciones si la función de utilidad de la unión es mejor comparada con cada coalición por individual, además si el set de coaliciones abarca todos los usuarios de la partición, y por Pareto, es preferible, dada su función de utilidad, comparada con las particiones no coalicionadas.
 - b) La comparación se realiza basada en la siguiente función de utilidad de cada usuario CR:

$$v(S) = Q_{d,S} - C(Q_{f,S}) = (1 - Q_{m,S}) - C(Q_{f,S}) \quad (2.3)$$

Donde $C(Q_{f,S})$ es la función costo de probabilidad de falsa alarma dentro de la coalición S.

- c) En el juego de detección de colaboración propuesto, la utilidad de una coalición S es igual a la utilidad de cada usuario CR en la coalición.
- d) Las probabilidades de falsa alarma y de perder la detección de un PU de cualquier usuario CR que pertenece a la coalición están dadas por dichas probabilidades pero de la coalición.
- **Separación:** Se decide separar un set de coaliciones si la función de utilidad de cada coalición del set por individual es mejor que la unión de las coaliciones
- 3) **FASE 3:** Detección de la coalición, cada usuario CR reporta su bit de detección a cada cabeza de la coalición. La cabeza de cada coalición toma una decisión final sobre la presencia o no de un usuario primario usando una regla OR

4. Etapa de compartición y movilidad: Selección de tecnología por Handoff mediante el parámetro RSSI

Debido a que la tecnología LTE es compatible con las especificaciones 3rd Generation Partnership Project (3GPP) LTE [9] y la tecnología WiFi es compatible con las especificaciones IEEE 802.11 [10], y dado que el trabajo se basó exclusivamente en métodos y análisis de la capa PHY CR, se investigó y analizó un parámetro físico característico común en ambas tecnologías, para poder considerarlo como parámetro de decisión para la movilidad de CR entre tecnologías. Este parámetro físico es el Received Signal Strength Indication (RSSI), dado que provee información acerca del nivel de potencia que está siendo recibido por la antena el cual decrementa cuando el usuario se aleja del actual punto de acceso a la red. Este criterio permite determinar el momento en el que se hace necesario realizar un cambio de canal, tal como lo explican Nahas, Mjalled y Ghantous [11].

El esquema seguido para realizar el Handoff se muestra en la Figura N° 2.

III. IMPLEMENTACIÓN

La implementación de los algoritmos y métodos analizados, se la realizó mediante un desarrollo para Network Simulator 3 (NS-3), y su resultado, fue un simulador CR para las redes móviles LTE y WiFi, siguiendo las mejores prácticas de desarrollo de software presentadas por Spillner, Linz y Schaefer [12]. El desarrollo se encuentra estructurado en 5 componentes principales:

- Módulo Cognitive
- Módulo LTE
- Módulo WiFi
- Programa principal para generación de simulación LTE-WIFI.CC
- Script para inicialización de parámetros Lte-Wifi.sh

1. Módulo Cognitive

Contiene los desarrollos que permiten realizar la etapa de detección del ciclo CR debido a que está implementado el método de detección SVD especificado en el inciso II.

2. Módulo LTE

Contiene todos los parámetros principales de la tecnología LTE, adaptada a las funciones cognitivas del desarrollo. Aquí se encuentra implementado el algoritmo de teoría de juego de coaliciones encargado de la decisión de uso del espectro, especificado en el inciso II.

3. Módulo WiFi

Contiene todos los parámetros principales de la tecnología WiFi, adaptada a las funciones cognitivas del desarrollo

4. Programa principal para la generación de simulación

Código principal donde se encuentran estructurados todos los parámetros técnicos básicos de la simulación

5. Script para inicialización de parámetros

Este script permite modificar los parámetros técnicos básicos de simulación, además para iniciar la simulación en el terminal se debe ejecutar este script.

Los parámetros técnicos modificables en el desarrollo con sus respectivos umbrales, se especifican en la Tabla N° 2

Tabla N°2. Tabla con umbrales de parámetros técnicos modificables

Parámetros	Valor mínimo	Valor máximo
Frecuencia LTE	729 MHz	2170 Mhz
Frecuencia WiFi	2400 MHz	5875 Mhz
Ancho de banda LTE	10 MHz	60 Mhz
Ancho de banda WiFi	20 MHz	150 Mhz
Celdas eNB	0	100
CR LTE UE	0	100
CR WiFi UE	0	100
CR UE dual	0	100
LTE UE primario	0	100
Rango de cobertura AP	200 m	500 m
Rango de cobertura eNB	350 m	5000 m
Tiempo de simulación	0	120 h
Número de muestras	1000	30000
Modelo de Movilidad	Random Waypoint	
Modelo de Propagación	Range Propagation Loss	

Los módulos desarrollados en NS-3 interactúan entre sí en un entorno cognitivo, como se muestra en el diagrama de interacción de los módulos implementados en la Figura N° 3

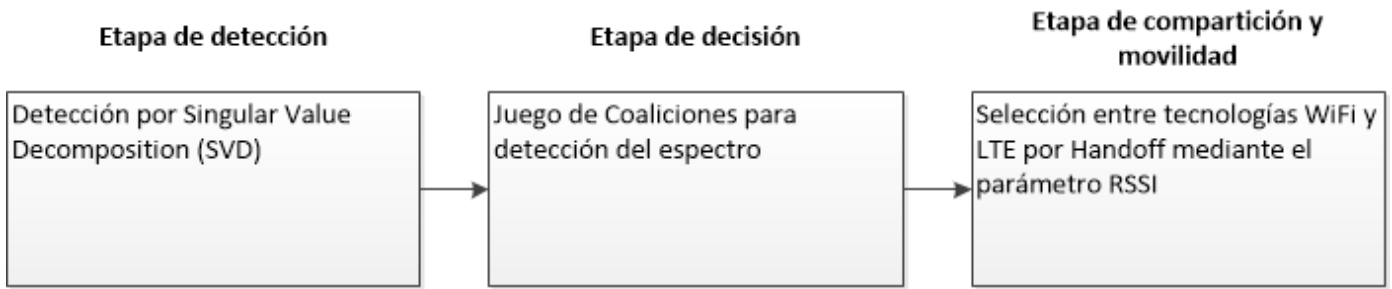


Fig. N°1. Diagrama de flujo para la metodología usada

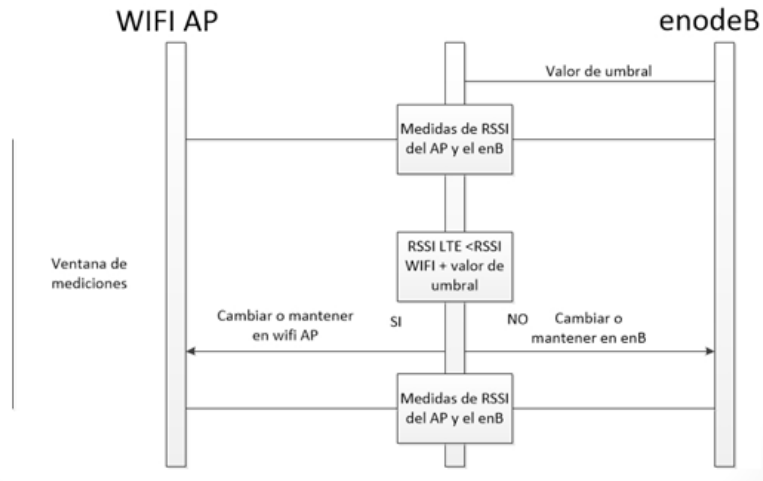


Fig. N°2. Esquema de Handoff entre tecnologías

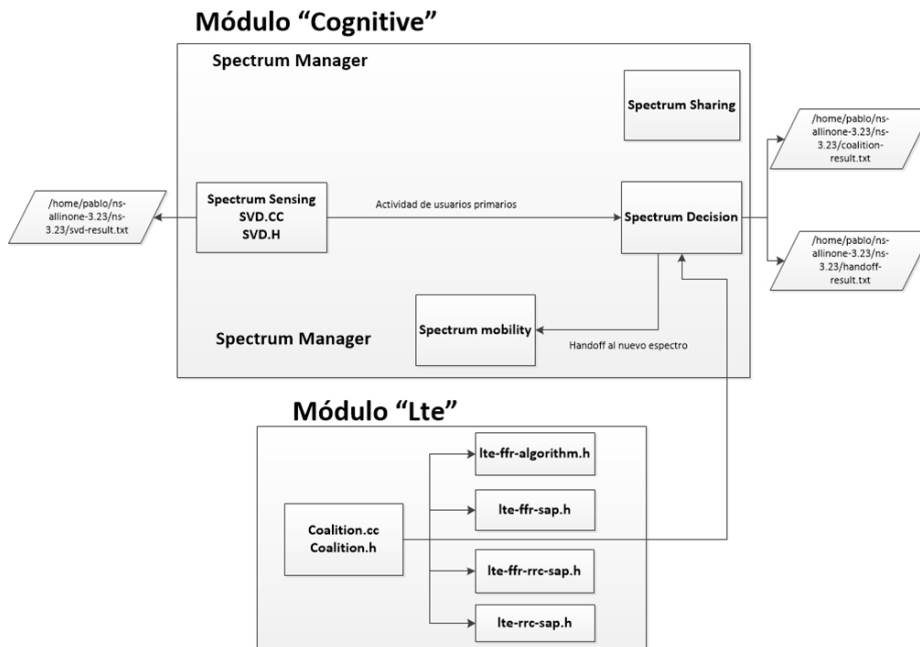


Fig. N°3. Diagrama de interacción de los módulos implementados

IV. RESULTADOS

Los datos del algoritmo SVD según el tipo de usuario CR se encuentran en un archivo .txt, es una salida al finalizar la simulación, que tiene la estructura mostrada en la Figura N° 4.

Svd-result.txt

Time	Type of User	Detection (0 not detect, 1 detect)	Primary User
0.220000	CR-Lte-0	1	EU-Lte-5
1.105.523	CR-Dual-8	1	EU-Lte-8
1.994.839	CR-Wifi-0	1	EU-Wifi-5
2.488.930	CR-Wifi-0	1	EU-Wifi-5
2.921.562	CR-Dual-15	1	EU-Lte-9
3.266.540	CR-Lte-2	1	EU-Lte-3
3.538.929	CR-Lte-2	0	EU-Lte-4
3.958.482	CR-Lte-2	1	EU-Lte-8
4.749.278	CR-Dual-6	1	EU-Wifi-1
4.964.039	CR-Dual-17	1	EU-Wifi-3
5.211.961	CR-Dual-4	1	EU-Lte-1
5.536.223	CR-Dual-16	1	EU-Lte-8
6.354.093	CR-Lte-3	1	EU-Lte-7
7.085.685	CR-Lte-4	1	EU-Lte-6
7.628.540	CR-Dual-0	1	EU-Wifi-1
7.889.481	CR-Dual-4	1	EU-Wifi-7
8.159.196	CR-Lte-3	0	EU-Lte-7
8.872.944	CR-Wifi-0	0	EU-Wifi-4
9.249.632	CR-Dual-3	0	EU-Wifi-3
9.888.745	CR-Wifi-3	1	EU-Wifi-9
10.661.858	CR-Dual-14	1	EU-Wifi-2
11.550.767	CR-Lte-3	1	EU-Lte-2
12.237.540	CR-Dual-11	0	EU-Lte-5
12.619.910	CR-Dual-12	1	EU-Lte-4
13.548.115	CR-Lte-0	1	EU-Lte-1
13.762.521	CR-Wifi-2	0	EU-Wifi-8
14.207.478	CR-Wifi-4	1	EU-Wifi-3

Probability of Detect: 0.777778 Probability of false Detect: 0.222222

Fig. N°4. Estructura de archivo de salida “SVD”

Los datos del algoritmo de coalición se encuentran en un archivo .txt, es una salida al finalizar la simulación, que tiene la estructura mostrada en la Figura N° 5.

Coalition-result.txt

Time	Coalition	Head of Coalition	Detection of the coalition	Technology
0.272243	CR-Dual-9,CR-Dual-1,CR-Dual-16,CR-Dual-17,CR-Dual-13,CR-Dual-18,CR-Dual-15,CR-Dual-8,CR-Dual-4,CR-Dual-8,	CR-Dual-15	0	lte
1.497.116	CR-Dual-4,CR-Dual-11,CR-Dual-12,	CR-Dual-11	1	wifi
2.783.173	CR-Dual-6,CR-Dual-7,CR-Dual-11,CR-Dual-13,CR-Dual-16,CR-Dual-5,CR-Dual-16,CR-Dual-17,	CR-Dual-7	0	wifi
3.477.865	CR-Wifi-4,CR-Wifi-3,CR-Wifi-0,CR-Wifi-2,CR-Wifi-3,CR-Wifi-1,CR-Wifi-4,CR-Wifi-3,CR-Wifi-3,CR-Wifi-1,	CR-Wifi-1	1	wifi
3.058.317	CR-Dual-0,CR-Dual-18,CR-Dual-4,CR-Dual-3,CR-Dual-15,CR-Dual-6,	CR-Dual-15	0	wifi
3.769.874	CR-Wifi-0,CR-Wifi-2,CR-Wifi-4,CR-Wifi-3,	CR-Wifi-3	0	wifi
4.253.031	CR-Wifi-2,CR-Wifi-0,CR-Wifi-0,CR-Wifi-2,CR-Wifi-4,CR-Wifi-3,CR-Wifi-1,	CR-Wifi-3	1	wifi
4.524.573	CR-Wifi-4,CR-Wifi-3,CR-Wifi-4,CR-Wifi-0,CR-Wifi-4,CR-Wifi-4,CR-Wifi-3,CR-Wifi-2,CR-Wifi-4,CR-Wifi-0,	CR-Wifi-4	1	wifi
5.137.489	CR-Dual-17,CR-Dual-10,CR-Dual-3,CR-Dual-16,CR-Dual-7,CR-Dual-16,CR-Dual-15,CR-Dual-11,	CR-Dual-16	0	wifi
5.011.231	CR-Lte-0,CR-Lte-2,CR-Lte-1,CR-Lte-1,CR-Lte-3,	CR-Lte-0	1	lte
6.050.778	CR-Dual-15,CR-Dual-12,CR-Dual-7,	CR-Dual-12	0	wifi
6.463.371	CR-Wifi-4,CR-Wifi-3,CR-Wifi-4,CR-Wifi-2,CR-Wifi-4,CR-Wifi-4,CR-Wifi-4,CR-Wifi-2,	CR-Wifi-4	1	wifi

Fig. N°5. Estructura de archivo de salida “coalition”

Los datos de handoff entre tecnología LTE y WiFi o viceversa se encuentran en un archivo .txt. El handoff se realiza de acuerdo al RSSI, la salida al finalizar la simulación, tiene la estructura mostrada en la Figura N° 6.

Handoff-result.txt

Time	User that makes the Handoff	Coalition to which belongs	From,RSSI	To,RSSI	Previous Channel	Current Channel
0.220000	CR-Dual-18	CR-Dual-9,CR-Dual-1,CR-Dual-16,CR-Dual-17,CR-Dual-13,CR-Dual-18,CR-Dual-15,CR-Dual-8,CR-Dual-4,CR-Dual-8,	Lte,33.746 ->	Wifi,36.123	lte	wifi
1.105.523	CR-Dual-4	CR-Dual-4,CR-Dual-11,CR-Dual-12,	Lte,29.050 ->	Wifi,34.868	lte	wifi

Fig. N°6. Estructura de archivo de salida “handoff”

Para poder generar datos de los algoritmos implementados en cualquier de las etapas se debe seguir el flujo de trabajo mostrado en la Figura N° 7, usado por Palacios, Castro, Azurdia y Estevez en [13], para realizar el análisis de comparación entre el algoritmo de detección SVD y algoritmos de detección teóricos.

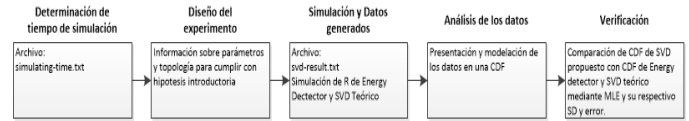


Fig. N°7. Diagrama de flujo para análisis de datos y resultados

Además, la topología que se debe utilizar para que los resultados sean confiables y acorde a la investigación realizada, se muestra como la red propuesta para simulaciones en la Figura N° 8.

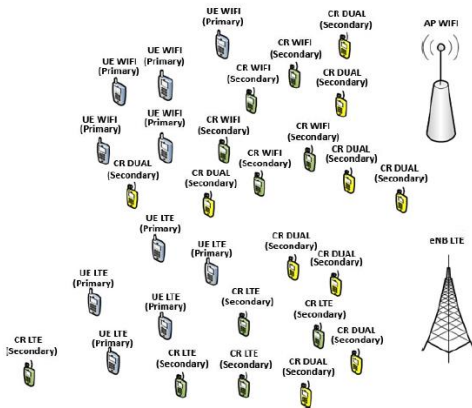


Fig. N°8. Topología de red propuesta para simulaciones

V. CONCLUSIONES

1. Dada la completa implementación del sistema CR en las tecnologías WiFi y LTE, este simulador se convierte en una potente herramienta no sólo para cumplir los objetivos del trabajo en cuanto a una única etapa del ciclo CR, sino para pruebas en cualquier otra etapa del sistema, además investigaciones posteriores en cuanto a rendimientos de la red, ya que nos permite variar los parámetros de entrada del simulador en rangos muy usados en las redes comerciales.

2. Las cuatro etapas de un sistema CR, detección, decisión, compartición y movilidad, fueron implementadas en los módulos cognitive, WiFi y LTE modificados, adaptados y compatibles con el software NS-3.

3. Los escenarios de simulación que se pueden diseñar tienen un alto grado de versatilidad, por lo que no sólo permiten comparar los algoritmos base aplicados en el simulador, sino que se pueden crear módulos para pruebas con otros algoritmos teóricos y poder compararlos mediante un correcto análisis estadístico y matemático.

4. El simulador fue desarrollado de tal manera que no requiera de una alta capacidad o conocimientos de programación computacional por parte del usuario, ya que contiene scripts que permiten modificar parámetros de manera sencilla, además de tener su propio manual de uso.

VI. REFERENCIAS

- [1] J. H. Aguilar Rentería and A. Navarro Cadavid, "Radio cognitiva—estado del arte", *Sistemas y Telemática*, vol. 9, no. 16, pp. 31–53, 2011.
- [2] J. Mitola, "Cognitive Radio Architecture Evolution," in *Proceedings of the IEEE*, vol. 97, no. 4, pp. 626-641, April 2009.
- [3] I. M. Kalil, "Cognitive radio the ieee 802.22 standard," *Integrated Communication Systems*, vol. 15, 2011.
- [4] D. B. Cabric, *Cognitive Radios: Systems Design Perspective*. PhD thesis, University of California, Berkeley, 2007.
- [5] M. F. Fahim and M. S. Raeen, "Svd detection for cognitive radio network based on average of maximum-minimum of the icdf," *International Journal of Advanced Computer Research*, vol. 2, Sep. 2012.
- [6] W. Saad, Z. Han, M. Debbah, A. Hjørungnes, and T. Basar, "Coalitional games for distributed collaborative spectrum sensing in cognitive radio networks," in *IEE INFOCOM 2009*, pp. 2114–2122, April 2009.
- [7] T. Wang, L. Song, Z. Han, and W. Saad, "Overlapping coalitional games for collaborative sensing in cognitive radio networks," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 4118–4123, Apr. 2013.
- [8] Q. Ni, R. Zhu, Z. Wu, Y. Sun, L. Zhou, and B. Zhou, "Spectrum allocation based on game theory in cognitive radio networks.," *JNW*, vol. 8, no. 3, pp. 712–722, 2013.
- [9] 3rd Generation Partnership Project, "Long term evolution overview." <http://www.3gpp.org/technologies/keywords-acronyms/98-lte>. Visitado: 2016-09-12.
- [10] I. S. Association et al., "Ieee 802.11TM-2012," standards. ieee.org/about/get/802/802.11.html. Visitado: 2016-09-13.
- [11] M. Nahas, M. Mjalled, Z. Zohbi, Z. Merhi, and M. Ghantous, "Enhancing lte – wifi interoperability using context aware criteria for handover decision," in *2013 25th International Conference on Microelectronics (ICM)*, pp. 1–4, Dec 2013.
- [12] A. Spillner, T. Linz, and H. Schaefer, *Software testing foundations: a study guide for the certified tester exam*. Rocky Nook, Inc., 2014.
- [13] P. Palacios, A. Castro, C. Azurdia-Meza and C. Estevez, "SVD detection analysis in cognitive mobile radio networks," *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)*, Milan, 2017, pp. 222-22

Scrum y Django para el desarrollo rápido de un sistema web para el control de mantenimientos preventivos y correctivos de bienes del cuerpo de bomberos del GADM de Santo Domingo

Silva, Hugo., Ocampo, Willian., Ulloa, Luis., Salazar, Diego., Córdova, Rodolfo., Andrade, Milton.
 hfsilvaa@pucesd.edu.ec, opwj@pucesd.edu.ec, umlj@pucesd.edu.ec, sadr@pucesd.edu.ec, cgrs@pucesd.edu.ec,
 asmt@pucesd.edu.ec

Pontificia Universidad Católica del Ecuador - Sede Santo Domingo – Ecuador

Resumen—Mediante la implementación de un sistema informático en un entorno web, se realiza una administración adecuada de los registros de los bienes y se mejora el control de los mantenimientos que lleva a cabo el Departamento de Logística del cuerpo de bomberos de Santo Domingo. El software SIS-CO (Sistema para el Control y Mantenimiento de los Bienes) tiene seis módulos: administración del sistema que permite manejar todos los parámetros y configuración global de la aplicación; personal, proveedores y bienes que son módulos que sirven para registrar los bienes con sus respectivas características; custodia que permite registrar la persona a cargo del bien y mantenimiento para cada bien registrado. En el presente documento se expone todas las herramientas y técnicas empleadas en el desarrollo del sistema SIS-CO. Se empleó el marco de trabajo Scrum para gestionar el proyecto, el patrón de diseño MVC, PostgreSQL 9.6 como sistema gestor de base de datos, el lenguaje de programación Python 3.5, el framework de desarrollo Django 1.8.18 LTS para la codificación del sistema, HTML5 y Bootstrap 3.3.6 (CSS3). Con la implementación del software se automatizó el proceso de seguimiento y control de mantenimientos de los bienes informáticos.

Palabras clave: Django, SCRUM, PostgreSQL, Bootstrap, software libre.

I. INTRODUCCIÓN

Es una investigación que tiene por objetivo controlar el mantenimiento de los bienes que administra el Cuerpo de Bomberos de Santo Domingo a través de un software que monitoree la periodicidad para realizar dichos mantenimientos. Los datos se obtienen de los documentos donde se registran los mantenimientos y en otros donde se lleva un inventario de los mismos.

Esta bibliografía se centra en las nuevas tendencias tanto en la selección de herramientas como la construcción del software, para así poder tener un producto robusto y escalable.

II. DESARROLLO

5. Marco de Trabajo

En la gestión de proyectos existe tres dimensiones fijas: alcance, costo y tiempo (triple restricción); con Scrum se

mantiene fijo el costo y tiempo mientras el alcance se desarrolla por Sprint.

Scrum se basa en la teoría de control de procesos empírica o empirismo [1]. El empirismo asegura que el conocimiento procede de la experiencia y de tomar decisiones basándose en lo que se conoce. Scrum emplea un enfoque iterativo e incremental.

Roles:

a. Roles

Product Owner: [1] es el responsable de gestionar la Lista del Producto desde la visión del negocio.

Scrum Master: [1] es el responsable de que Scrum se entienda y se adopte, garantizando que se respete los principios del marco de trabajo.

Equipo de Desarrollo: encargados de la construcción el producto.

Artefactos: [1] representan trabajo o valor en diversas formas que son útiles para proporcionar transparencia y oportunidades para la inspección y adaptación.

b. Artefactos

Product Backlog: [1] Es una lista ordenada de todo lo que podría ser necesario en el producto y es la única fuente de requisitos para cualquier cambio a realizarse en el producto., y el propietario de este es el Product Owner. Una Lista de Producto nunca está completa. El desarrollo más temprano de la misma solo refleja los requisitos conocidos y mejor entendidos al principio.

Sprint Backlog [1]: es el conjunto de elementos de la Lista de Producto seleccionados para el Sprint, más un plan para entregar el Incremento de producto y conseguir el Objetivo del Sprint., es propiedad del equipo.

Incremento: [2] la parte de producto producida en un sprint, y tiene como característica el estar completamente terminada y operativa, en condiciones de ser entregada al cliente.

Burndown Chart: [2] Lo actualiza el equipo en el scrum diario, para comprobar si el ritmo de avance es e previsto, o se puede ver comprometida la entrega del sprint.

c. Eventos:

Sprint: [2] de esta manera se nombra a cada iteración de desarrollo.

Planificación del Sprint (Sprint Planning): En esta reunión se planifica el Sprint, y al finalizar esta se debe un

objetivo claro y compartido sobre el trabajo que hay que realizar para la siguiente iteración y con un Sprint Backlog apropiado.

La Reunión Diaria (Daily Meeting): [3] “es la situación donde se sincroniza el equipo, cada miembro habla acerca del estado que se encuentra el trabajo que está elaborando”.

Revisión del Sprint (Sprint Review): al terminar el Sprint, el equipo analiza el estado del trabajo con el Product Owner y con cualquier persona que pueda contribuir con información útil, no se debe invertir mucho tiempo, se la debe realizar informalmente; es la oportunidad de mejorar la manera en la que se está construyendo.

Retrospectiva del equipo: luego de la Review, [3] el equipo se agrupa para buscar la mejora continua en su trabajo y examinar aspectos que impiden mejorar el rendimiento, es el lugar para analizar cómo mejorar lo que estamos trabajando.

2 Aspectos Técnico del Desarrollo

En el desarrollo del sistema web, se usó software libre; en específico la distribución Ubuntu Server, junto con el lenguaje de programación Python, [4] porque es un lenguaje robusto, fácil de mantener cualquier tipo de software que se desarrolle, posee una gran comunidad, es escalable y la curva de aprendizaje es corta, esto permite que sea fácil de leer el código fuente porque tiene una sintaxis limpia y sencilla. El Framework Django [5] proporciona muchas características y ventajas como las seguridades tanto en el manejo de usuarios, sus permisos y sesiones, así como protección contra ataques de inyección de código SQL, ataques de sitio cruzado (CSRF), entre otras utilidades listas para usar. Para el almacenamiento de la información de la base de datos, se utilizó PostgreSQL, [6] dado que provee un enorme potencial por su tiempo que lleva en desarrollo y mantenimiento, posee varias extensiones que nos proporciona más y mejores funcionalidades, muy robusto, estable y seguro dado a que su robustez al manejar una gran cantidad de volúmenes de datos y transacciones; proporciona soporte para varios lenguajes de programación.

También se utilizó el LiClipse por su integración con PyDev para manejar proyectos Python y Django. Mientras que el Framework Bootstrap permite crear interfaces amigables al usuario, ReportLab en la generación de reportes en PDF.

El diseño final de la base de datos del proyecto está en la Figura 1.

El modelo arquitectónico lógico o la arquitectura lógica empleado para el desarrollo del software es el MVC [7] el cual divide el software para obtener un mejor desempeño del sistema, uno de los mejores arquitecturas para el desarrollo de páginas web es el uso de arquitecturas en capas.

El framework Django fue diseñado para promover el acoplamiento débil y separación estricta entre sus componentes, debido a esto sigue el patrón MVC mencionado antes, con una ligera variación: Diseño MTV, en el cual M es el Modelo, en otras palabras la capa de acceso a la base de datos; T es la capa de presentación de datos, el Template (plantilla); y V, la Vista, la capa que tiene la lógica del negocio y la que accede al Modelo.

El software es accesible solo en la intranet de la institución donde se hicieron las pruebas antes de ser puesto en producción, y trabaja como apoyo junto a las otras aplicaciones que ya tiene la institución; este software se encarga del seguimiento y control de mantenimientos de los bienes de la institución, de esta manera se mejora

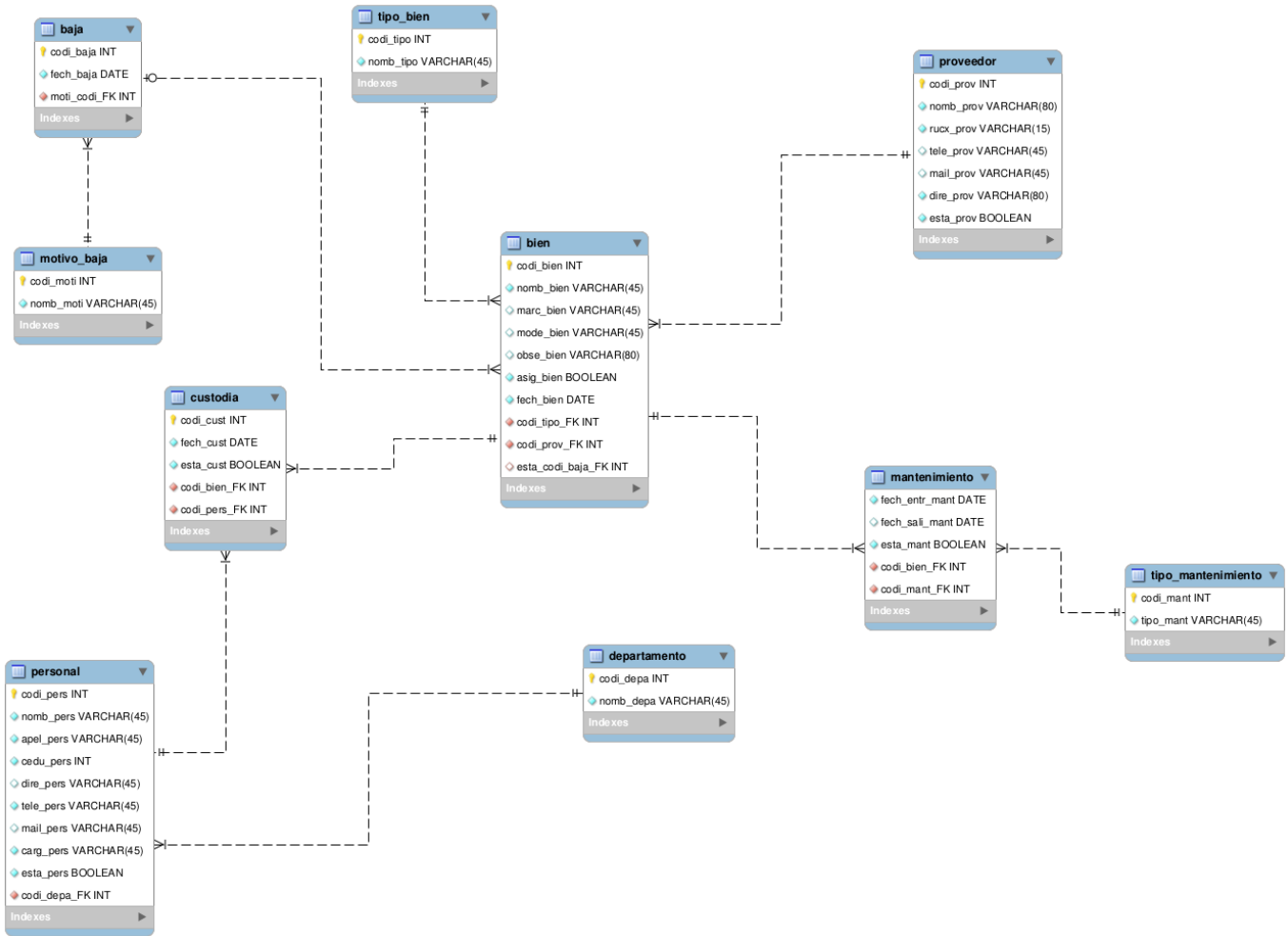


Fig. 1 Diseño Final de Base de Datos

la productividad [8] y dichos mantenimientos prolongan la vida útil del bien, todo

esto representa una inversión dentro de un periodo de tiempo mediano o largo que originará ganancias.

Para el desarrollo del software se realizó por sprints que dicta la metodología de desarrollo de software, en cada uno de los sprints se realiza y entrega un módulo al final, completamente funcional y listo para ser puesto en producción, y su respectivo diseño de bases de datos para el funcionamiento de dicho módulo. Al finalizar todos los sprints se tiene un diseño final de la base de datos, como se puede observar en la Figura 2, el framework genera un historial de los cambios realizados en los diseños de la base de datos.

Para el desarrollo y seguimiento del proyecto se utilizó el sistema de control de versiones Git que provee mantenimiento de versiones de aplicaciones cuando tienen muchos archivos en el código fuente, junto con un repositorio creado en github.com.

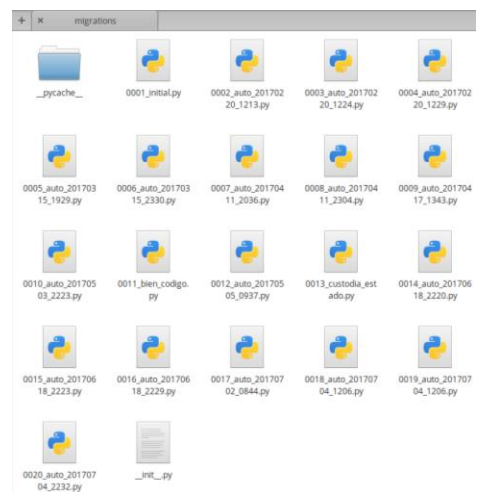


Fig. 2 Historial de Migraciones

Esto permite llevar un historial de los avances que se realizan en el código fuente del proyecto y de esta manera tener

backups del proyecto. El software que permite realizar estas sincronizaciones es GitKraken en GNU/Linux; es una herramienta de libre, con esta aplicación podemos sincronizar el código fuente con el repositorio fácilmente a través de una interfaz gráfica, también se puede acceder a las versiones

anteriores sincronizadas desde la misma aplicación, sin necesidad de acceder directamente a github.com; la Figura 3 es el historial del proyecto y reemplaza al Gráfico del Producto.

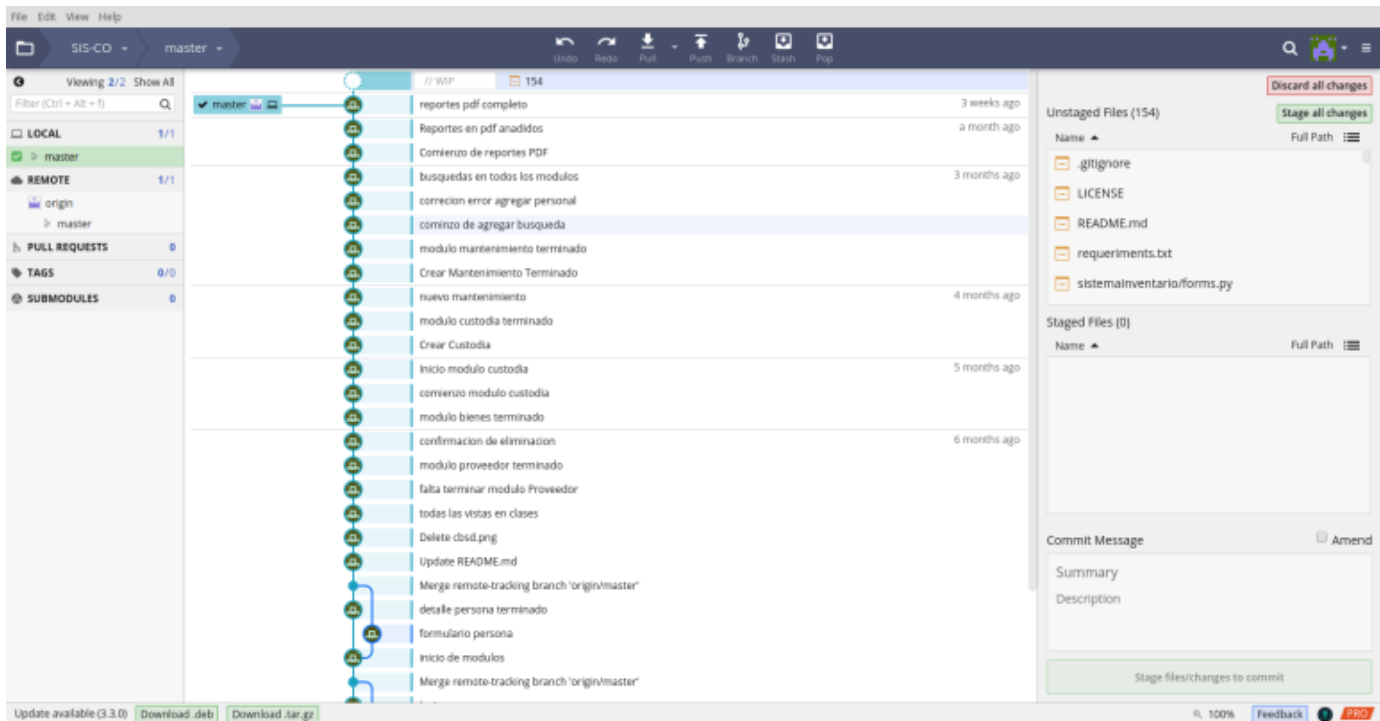


Fig. 3 Historial del proyecto (Gráfico del Producto)

Al final del desarrollo del sistema se tiene un diseño de bases de datos con todas las tablas de los diferentes módulos relacionadas. En cada Sprint se realizó un planteamiento de tablas con sus respectivas relaciones para que el módulo funcione correctamente, y a medida que crecía el número de tablas de la base de datos se hacía las respectivas relaciones para integrarla en un solo diseño, de esta manera se obtiene una sola base de datos del sistema.

Tener un buen diseño de base de datos mejora considerablemente la consistencia e integridad de los datos, se reduce la redundancia de la información y se evita que las búsquedas arrojen información incorrecta o errónea. Una base de datos es el cimiento de cualquier tipo de aplicación. Para el desarrollo de la aplicación se empleó Django que proporciona variadas características y ventajas entre ellas la seguridad. Django permite un desarrollo rápido de aplicaciones web, fomenta un diseño limpio, pragmático y la reutilización de código. La versión empleada es 1.8.18 LTS la cual tiene soporte hasta el año 2018, LTS por sus siglas Long Term Support, soporte de larga duración, también se utilizó Python esto permite que se migre fácilmente a futuras versiones.

La interfaz administrativa del sistema hace uso del sitio administrativo que provee Django, en el cual se debe registrar los modelos de la base de datos y configurar los parámetros necesarios para que entre en funcionamiento, como resultado

de esto se tiene un sitio totalmente independiente del proyecto, y se conserva la programación orientada a objetos que menciona la reutilización de código y la filosofía del framework, no repetir código.

Este sitio de administración es potente y se encuentra listo para ponerse en producción y se puede agregar, ver, editar y eliminar información de la base de datos también se registra todos los movimientos o cambios que se realice con el usuario, a esta interfaz solo pueden acceder quienes sean super-usuario y tengan los respectivos permisos; los usuarios del sistema no tendrán acceso a menos que se les otorgue los privilegios necesarios. También cuenta con un diseño predefinido que facilita al usuario su manejo y mejora la comunicación e interacción con el mismo, como se puede ver en la Figura 4.

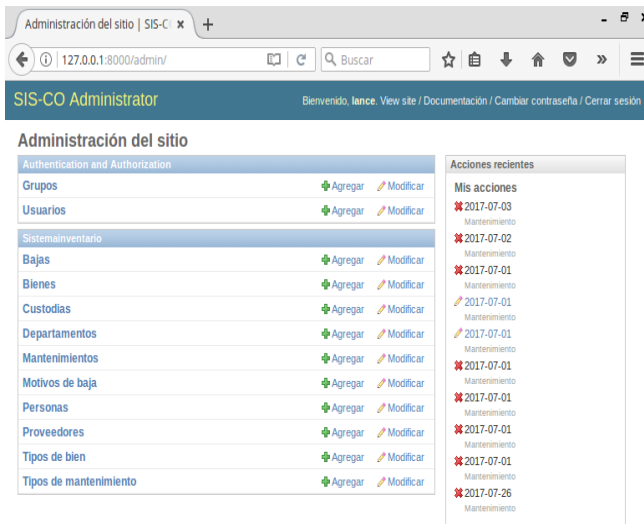


Fig. 4 Sitio Administrativo del Sistema

Para el diseño de las interfaces del sistema se utilizó el Framework Bootstrap una de las principales características es la rapidez para crear diseño web, también emplea el responsive web design (diseño web adaptable) por tal motivo todos los elementos creados se adaptan a la perfección a cualquier tipo de dispositivo y a sus diferentes tamaños de pantallas, no importa si se visualiza desde una computadora, una tablet, o un smartphone siempre se tiene una interfaz fácil de utilizar como se puede observar en la Figura 5.

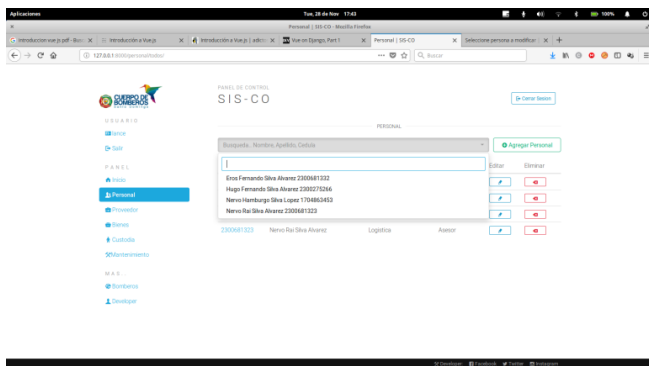


Fig. 5 Interfaz gráfica del usuario

2 Planificación

La primera reunión que se debe realizar es un Release Planning, esta se realizó con el Product Owner y las personas que conocen o tienen las características que tendrá el producto.

Luego se realiza el Sprint Planning con el Product Owner, se hizo una reunión en la cual se discutió y analizó las historias de usuarios para mejorarlas en el Sprint 0, de esta manera fue posible la elaboración del Product

Backlog, ver tabla 1; la duración de cada Sprint se definió de cuatro semanas para desarrollar las historias de usuario asignadas a cada uno de los Sprints. Se Ordenó las funcionalidades en base a la prioridad que propuso el Product Owner.

Tabla 1 Product Backlog

Product Backlog			
#	Historia	Prioridad	Estimación
1	Acceso de usuarios	100	13
2	Ingreso de empleado	100	8
3	Buscar y editar datos del empleado	95	13
4	Eliminar empleado	95	13
5	Acceso de administrador	95	13
6	Ingreso de usuario	95	8
7	Agregar y editar permisos de usuarios	95	5
8	Eliminar usuario	90	5
9	Acceso total al módulo personal	90	13
10	Ingreso de proveedor	90	8
11	Buscar y editar datos del proveedor	85	13
12	Eliminar proveedor	85	13
13	Acceso total al módulo proveedor	85	13
14	Ingreso de bienes	85	13
15	Buscar y editar datos del bien	80	13
16	Eliminar bienes	75	13
17	Acceso total al módulo bienes	75	5
18	Ingreso de custodias	75	8
19	Buscar y editar datos de la custodia	75	13
20	Eliminar custodias	70	13
21	Acceso total al módulo custodia	70	8
22	Ingreso de mantenimientos	70	21
23	Buscar y editar datos de mantenimiento	70	8
24	Mantenimiento realizado	65	8
25	Acceso total al módulo mantenimiento	65	5

Para la estimación de las historias de usuario se empleó la métrica puntos de historia y la técnica serie Fibonacci la cual consiste en dar la valoración midiendo la complejidad de la historia, particularmente en este caso se acortó la serie del uno al veinte y uno, donde 1, 2 y 3 es poco compleja; 5 y 8 medianamente compleja; 13 y 21 compleja.

La planificación anteriormente descrita se la realizó antes de iniciar cada Sprint (Sprint Planning) y es donde se selecciona las historias de usuario que serán parte de cada Sprint Backlog ver tabla 2, manteniendo la prioridad que considere el Product Owner.

Tabla 2 Sprint Backlog 0

N° Historia	Estimación	Categoría	Tarea	Estimación
1	13	Desarrollo	Configuración del O.S. para instalar Django 1.8	1
		Desarrollo	Instalación y configuración de PostgreSQL 9.6	1
		Diseño	Crear diseño de la base de datos de la historia	3
		Diseño	Validar diseño de la base de datos	4
		Desarrollo	Migrar modelo a la base de datos	1
		Desarrollo	Crear diseño web para el acceso de usuarios	2
		Desarrollo	Validar diseño y acceso de usuarios	1
2	8	Diseño	Crear diseño de la base de datos de la historia	2
		Diseño	Validar diseño de la base de datos	2
		Desarrollo	Crear diseño web para el formulario	2
		Desarrollo	Validar diseño e ingreso de datos en formulario	2
3	13	Diseño	Crear diseño de la base de datos de la historia	2
		Diseño	Validar diseño de la base de datos	2
		Desarrollo	Crear diseño web para listar y la buscar	3
		Desarrollo	Crear diseño web para el formulario	3
		Desarrollo	Validar diseño de listar, buscar y editar datos	3
4	13	Diseño	Crear diseño de la base de datos de la historia	1
		Diseño	Validar diseño de la base de datos	2
		Desarrollo	Crear diseño web para listar y eliminar	3
		Desarrollo	Validar diseño de listar y eliminar datos	4
		Desarrollo	Agregar avance al sistema de control de versión	3

3 Daily Meeting

En el Daily Meeting exponía lo que se ha hecho desde la última reunión, los inconvenientes que se han presentado o algún riesgo que se deba revisar y lo que se planea realizar para la siguiente reunión, se hizo uso de la herramienta de gestión de tareas Trello. (<https://trello.com/b/n8LhGFNM/sis-co>)

Aquí se actualizaba el trabajo y avance día a día, a su vez se reflejaba en los cuadros de avance diario (Burn Down Chart), ver Figura 6.

En la ejecución del Sprint (Sprint Execution) No se puede rebajar la calidad de los objetivos del Sprint y todo software que se entrega debe estar listo para ponerse en producción; los requerimientos que no se han desarrollado deben ser registrados o documentados para incluirlas en el Product Backlog y solo se realizan más adelante con la revisión y autorización del Product Owner.

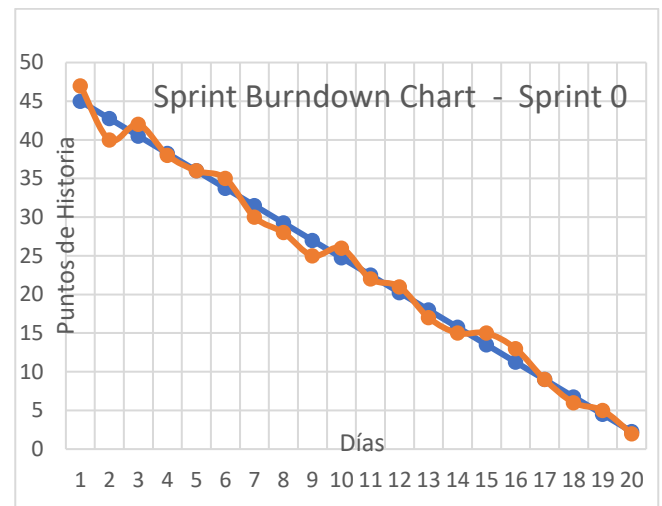


Fig. 6 Burn Down Chart (Gráfico de Avance Diario)

En el Refinement; a pesar de haber hecho un refinamiento en el Sprint Planning; también se puede realizar algo en el desarrollo de software porque en esta parte pueden aparecer inquietudes adicionales.

4 Sprint Review

La revisión del Sprint (Sprint Review) el Product Owner analiza el estado actual del Product Backlog y realiza una estimación de fechas basándose en el progreso actual; todo el grupo hace una colaboración para la continuación a modo de revisión.

5 Sprint Retrospective

La retrospectiva del Sprint (Sprint Retrospective) permite al equipo un autoanálisis para revisar el desempeño del proceso SCRUM, de esta manera crear un plan de mejoras para el siguiente Sprint y el propósito de esto es: revisar todo respecto a personas, relaciones, procesos, herramientas utilizadas, identificar y ordenar los temas principales que se presentaron, como se hizo el trabajo y si se realizaron todos los eventos en el tiempo definido.

III. RESULTADOS EXPERIMENTALES

En la Figura 7 se muestra el login de administrador o super-usuario. Está destinado al jefe del departamento de logística. Esta interfaz administrativa está pensada para administrar y controlar el funcionamiento sistema. Solo puede acceder el usuario que tiene los permisos necesarios

Fig. 7 Login Sitio Administrativo

Después de haberse identificado el usuario podrá acceder a ver y revisar en detalle todos los módulos del sistema sin restricciones como se muestra en la Figura 4, indicada anteriormente. El administrador, podrá crear otro administrador o a su vez asignar los respectivos permisos a los usuarios que desee darle acceso a esta interfaz.

Los usuarios sin permisos podrán ingresar al sitio sin privilegios, como se indica el login en la Figura 8.

Fig. 8 Login Usuario

Aquí podrán manipular e ingresar solo la información respectiva a sus permisos como usuarios, no se contempla la eliminación física de la información.

Luego de haberse identificado el usuario correctamente con su respectivo nombre de usuario y contraseña se redirige a la ventana principal del Panel de Control para que haga uso de los módulos que tiene habilitados para poder interactuar, ver la Figura 9. Si no hay notificaciones de reportes de mantenimientos por realizar, solo se muestra la página de bienvenida como en este caso.

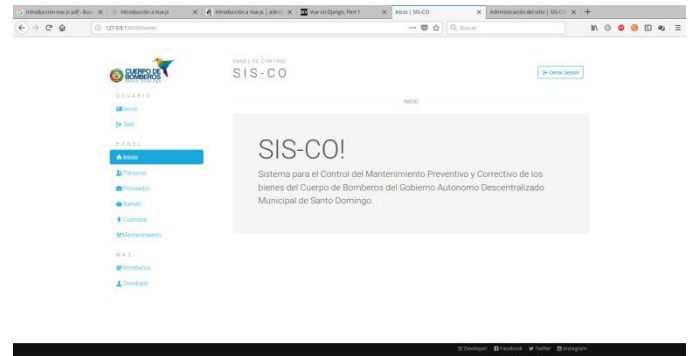


Fig. 9 Página de Inicio

El módulo de personal permite agregar personas que laboran en la institución, para posteriormente asignarles bienes que tienen a su custodia.

La administración de proveedores se realiza en la interfaz de proveedor, donde se agregan estos para luego poder registrar los bienes que posee la institución.

Para el manejo de los bienes es necesario haber registrado el proveedor de dicho bien. Todos los módulos se controlan a través de estados que permiten filtrar los registros de acuerdo a dicho estado; en el estado activo todos los registros se pueden manipular y en el estado desactivado no se permite su modificación o acceso.

El módulo de custodia se encarga de enlazar al personal con sus respectivos bienes que tiene a cargo. Se desarrolló un acceso para todo el personal que tiene en su custodia bienes con el fin de que puedan notificar las incidencias, de esta forma se automatiza y gestiona de mejor manera los registros de los mantenimientos así como sus notificaciones; puede verse esto en la Figura 10.

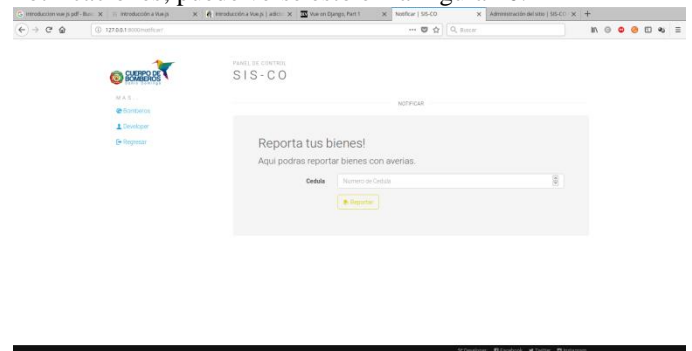


Fig. 10 Notificaciones de incidencias

En esta sección solo accede a reportar bienes con incidencias el personal registrado, y si en algún caso no tiene bienes asignados no podrá registrar ninguna notificación, todas las notificaciones son registradas junto con la persona que la realiza, y las anomalías que presenta el bien. La interfaz de notificaciones permite registrar incidencias de los bienes a su custodia. Los datos registrados no pueden ser eliminados por motivos de auditorías informáticas.

Cada módulo contiene reportes en formato pdf que ayudan en la toma de decisiones, la generación de reportes fue implementada con la librería ReportLab, escrita en Python. Esta librería encaja perfectamente con el Framework Django, la misma que tiene buena

documentación para su uso, y facilita el manejo de imágenes. Django permitió implementar seguridades en las conexiones a la base de datos encriptándolas, el manejo de sesiones de usuarios, se realiza de tal manera que se crea una tabla hash a partir de la misma sesión, por tal motivo en un supuesto ataque solo podrán ver el código hash, así se mantiene seguro el nombre de usuario y su contraseña; las contraseñas de los usuarios se encriptan.

Los formularios necesitan obligatoriamente un token para guardar la información del mismo en la base de datos, de esta forma se evita los ataques de sitios cruzados y es un nivel extra de seguridad, el framework, también provee utilidades para el manejo y validación de dichos formularios.

Quedan descartados los ataques de inyección de código SQL, por el hecho que Django provee un ORM (mapeo objeto-relacional) que accede a la base de datos, en la programación de las vistas y modelos que representan al diseño de la base de datos, se usan clases y objetos Python para realizar las consultas de los datos.

IV. CONCLUSIONES

1. Con la implementación del sistema se optimizaron los procesos referentes al mantenimiento de bienes como en el registro y el seguimiento de los mismos; esto reduce los tiempos de repuesta invertidos en los procesos del Departamento de Logística; aumenta la productividad en la institución y permite un mejor control sobre el inventario y el seguimiento de los bienes.

2. De acuerdo al análisis realizado previamente y la experiencia adquirida en el desarrollo del presente trabajo de investigación se usa SCRUM porque gestiona adecuadamente un proyecto que tiene requerimientos dinámicos y en los cuales se requiera un rápido despliegue, esto es posible gracias al manejo de sprints cortos de desarrollo que dan como resultados entregables totalmente funcionales del producto final.

3. Las herramientas utilizadas en el desarrollo de software permiten una elaboración rápida de aplicaciones, gracias a la integración que ofrece el framework Django con PostgreSQL; el lenguaje Python es potente en comparación con otros lenguajes de la misma categoría, permite escribir código fácilmente legible y mantenible.

4. Se construyó rápidamente la aplicación con el uso de Django, debido que provee compatibilidad con varios sistemas gestores de bases de datos; mejora significativamente la velocidad del desarrollo de software con un equipo mínimo de desarrolladores.

REFERENCIAS

- [1] Schwaber, K., & Sutherland, J. (2016). La Guía Definitiva de Scrum: Las Reglas del Juego. pp. 3 – 12.
- [2] Menzinsky, A., López, G., & Palacio, J. (2016). Scrum Manager. pp. 25 - 26, 45.
- [3] Álvarez, A., de las Heras, R., & Lasa, C. (2012). Métodos Ágiles y Scrum. Madrid: EDICIONES ANAYA MULTIMEDIA, pp 41.
- [4] Joyanes, L. (2013). Fundamentos generales de programación. México: McGraw Hill, pp 14 – 16.
- [5] Fernández, A. (2013). Python 3 al descubierto (Segundaed.). México: AlfaomegaGrupo Editor, pp 215.
- [6] The PostgreSQL Global Development Group. (2017). About PostgreSQL. Obtenido de <https://www.postgresql.org/about/>
- [7] López, M., Vara, J., Verde, J., Sánchez, D., Jiménez, J., & de Castro, V. (2014). Desarrollo Web en Entorno Servidor. Madrid: RA-MA, S.A. Editorial y Publicaciones, pp 96.
- [8] Bravo, M. (2011). Contabilidad General (Décima ed.). Quito: Editora ESCOBAR impresores, pp 170 – 1

Generador automático de aplicaciones web e interfaces de usuarios con funcionalidad responsiva en el lenguaje Python

Luis Herrera-Izquierdo ¹, Xavier Quiñonez Ku ¹, Juan Casierra Cavada ¹

luis.herrera@pucese.edu.ec, xavier.quinonez@pucese.edu.ec, juan.casierrac@pucese.edu.ec

¹Pontificia Universidad Católica del Ecuador - Sede Esmeraldas – Ecuador

Resumen — En la ingeniería de software, es importante la optimización del proceso de desarrollo y el aseguramiento de calidad. El objetivo de la optimización es evitar la realización de actividades repetitivas, que produzcan pérdida de tiempo y retraso en los periodos de entrega. Con el creciente uso de internet, las aplicaciones web, deben tener más funcionalidades que permitan satisfacer las necesidades de los clientes, por esta razón; es importante el desarrollo de aplicaciones con buenas prácticas fáciles de escalar. En esta investigación, se presenta un generador con funcionalidad responsiva, basado en buenas prácticas que permite crear aplicaciones web bases e implementar las respectivas operaciones CRUD mediante una interfaz amigable con el usuario. El generador está basado en el lenguaje de programación Python y el sistema gestor de bases de datos MySQL para el procesamiento de información, adicionalmente se hace uso de los frameworks Django y Bootstrap. Los resultados obtenidos del generador de código demostraron que se pueden generar aplicaciones web en periodos cortos de tiempo con funcionalidad responsiva, seguras y de alta calidad.

Palabras Clave – aplicaciones; web; interfaces; responsiva; django; framework; python; bootstrap.

I. INTRODUCCIÓN

Debido a que las capacidades de procesamiento de las computadoras han crecido exponencialmente al igual que las posibilidades de comunicación y transferencia de datos, resulta imprescindible desarrollar soluciones ágiles que se adapten al entorno de trabajo. En el mercado existen varios lenguajes de programación que permiten desarrollar soluciones de manera ágil, entre los cuales se pueden destacar: Java, Ruby, PHP, Python, Go, JavaScript entre otros. Las características principales que se toma en cuenta al momento de elegir un lenguaje de programación son [1]: productividad de desarrollo, calidad de software, portabilidad del programa, soporte de las librerías, componentes de integración, entre otras. Sin embargo, elegir un buen lenguaje de programación, no garantiza un producto de software de alta calidad, puesto que el desarrollo es una labor que requiere un alto conocimiento técnico del lenguaje de programación por parte del equipo de trabajo, para de esta manera implementar soluciones basadas en buenas prácticas.

La mayoría de los lenguajes de programación poseen un framework, que es un esquema (un esqueleto, un patrón) para el desarrollo y/o la implementación de una aplicación. Elegir un framework que facilite el desarrollo de la aplicación web, puede ser un elemento determinante de éxito o fracaso, por esta razón es importante conocer las ventajas y desventajas de los distintos frameworks disponibles antes de su elección [2].

Python es un lenguaje de programación popular debido a su simplicidad y la vez a su gran potencial, actualmente existen varios frameworks basados en este lenguaje, en la

siguiente figura se muestra un ranking de los frameworks más populares que está basado por el número de estrellas obtenidos del repositorio GitHub y en el número de preguntas realizadas en la plataforma Stack Overflow.

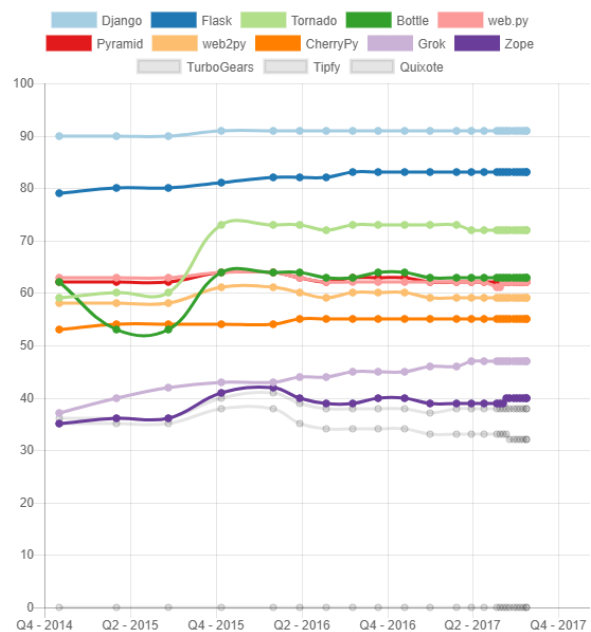


Fig.1. Ranking frameworks Python [5]

En la Fig. 1 se puede observar que el framework Django en los tres últimos años ha tenido una acogida positiva por los desarrolladores webs y está por encima de los demás frameworks basados en el lenguaje de programación Python.

En el mercado actual, existen frameworks orientados a la web, que poseen la funcionalidad de generar código como lo son: Ruby on Rails, Django o Yii2. Estos frameworks, facilitan la tarea de codificación a los desarrolladores.

En cuanto a la comparación entre Ruby on Rails y Django, ambos comparten las misma filosofía y son potentes pero es conveniente usar Django cuando el programador tiene un conocimiento técnico amplio del lenguaje de programación y es perfeccionista con los plazos de entrega de proyectos [3], en cuanto a Yii2, este es un framework que utiliza el patrón MVC, también es potente, sin embargo le permite al programador aplicar toda la lógica de negocio y acceso a la base de datos en la vista sin ninguna restricción, esto no es bueno cuando se quiere tener un proyecto de software de alta calidad y escalable [4].

En base a los análisis de los frameworks mencionados anteriormente, se requiere un generador que permita crear aplicaciones web en periodos cortos de tiempos con una estructura basada en buenas prácticas, por esta razón para esta investigación se eligió el framework Django.

Basados en el lenguaje de programación Python, el cual es un lenguaje de programación de uso general que se aplica a menudo en roles de scripting [6], en conjunto con el framework Django y Bootstrap. En la presente investigación se muestra un modelo generativo que permite la creación de aplicaciones web con interfaces responsivas, que se pueden ejecutar en teléfonos inteligentes, computadoras, televisores inteligentes y cualquier dispositivo que disponga de un navegador web, utilizando plantillas y tablas del sistema gestor de base de datos MySQL.

Además, se detallan las herramientas utilizadas para el desarrollo del generador de código, así como también su ejecución, basándose en un sistema de control de inventarios y para finalizar se realizó un análisis de los beneficios que se puede obtener al utilizar un generador de código y los trabajos a futuros que se pueden realizar.

II. MÉTODO

Durante las distintas etapas de desarrollo en los que hay operaciones repetitivas, resulta conveniente usar generadores de código que permitan automatizar esas tareas, de esta manera se reduce la probabilidad de cometer errores y por ende un ahorro en tiempo y dinero, [7] indica que en promedio, los desarrolladores de software gastan el 50% de su tiempo de programación en encontrar y corregir errores, al proyectar esta cifra en el costo total de desarrollo de software produce un costo de depuración que asciende a \$312 mil millones por año.

Cuando el software es difícil de crear o modificar, los programadores pasan la mayor parte del tiempo concentrándose en hacer que las cosas "funcionen", y menos tiempo para ayudar al usuario. Pero cuando un sistema es fácil de trabajar, los programadores pueden dedicar más tiempo a prestar ayuda al usuario y menos tiempo a centrarse en los detalles de la programación. Del mismo modo, cuanto más fácil es mantener una pieza de software, más fácil es para los programadores asegurarse de que el software sigue siendo útil [8].

Al momento de desarrollar aplicaciones de software aparte de crear soluciones eficientes que responden a las reglas de negocio también se debe tener en cuenta los aspectos de seguridad. Garantizar la seguridad es un proceso complejo y requiere conocimientos sobre diversos factores, como el rendimiento y la escalabilidad. Si una brecha o bug se pasa por alto durante el proceso de pruebas del software y se pone en producción, la explotación de esta brecha de seguridad podría afectar seriamente la continuidad del negocio [6].

El método propuesto usa técnicas de generación de código, la cual elimina el potencial de error humano en el proceso de configuración del proyecto web, esto es posible porque también se generan parámetros de configuración al momento de la generación del código fuente. En la solución planteada se utilizó el lenguaje de programación Python puesto que es un lenguaje de programación que está diseñado para ser legible, reutilizable y mantenible, mucho más que los lenguajes de scripting tradicionales [1], además es uno de

los lenguajes de programación más usados según el Spectrum's Top Ten Languages for 2016 [9], en la figura N° 2 se muestra la lista:











Language Rank	Types	Spectrum Ranking
1. C		100.0
2. Java		98.1
3. Python		98.0
4. C++		95.9
5. R		87.9
6. C#		86.7
7. PHP		82.8
8. JavaScript		82.2
9. Ruby		74.5
10. Go		71.9

Fig. 2. Spectrum's Top Ten Languages for 2016 [9]

Como ya se ha mencionado anteriormente, Python es un lenguaje de gran potencial en scripting, originalmente no está enfocado a la web así que para hacer esto posible se utilizó el framework Django, el cual está basado en el patrón: modelo, template, vista (MTV) y fue diseñado para que las tareas comunes de desarrollo de la web fueran rápidas y fáciles [10].

Una de las ventajas del framework Django, es que incentiva la reutilización mediante la estructuración de los proyectos en aplicaciones independientes que se comunican entre sí, además implementa varios mecanismos de seguridad lo que garantiza un nivel de confianza en las aplicaciones desarrolladas. Django cuenta con una gran comunidad de desarrolladores que constantemente están implementando funcionalidades y optimizando el framework, esto quiere decir que cada cierto periodo de tiempo se publican nuevas versiones. Para esta investigación se utilizó la versión de Django 1.10

Para el almacenamiento y procesamiento de información se utilizó el sistema gestor de base de datos (SGDB) MySQL, el cual es ampliamente conocido por su simplicidad y rendimiento. MySQL tiene como principal objetivo ser una base de datos fiable y eficiente. Ninguna característica es implementada en MySQL si antes no se tiene la certeza que funcionará con la mejor velocidad de respuesta y, por supuesto, sin causar problemas de estabilidad [11].

El acceso a la web mediante el uso de internet ha tenido un gran aumento en los últimos años [12], esto ha obligado a que las aplicaciones web tengan la capacidad de adaptarse a cualquier tipo de pantalla al momento de mostrar su contenido, es decir tener un diseño responsivo. Implementar la funcionalidad de responsividad a las aplicaciones web, es posible mediante la combinación del lenguaje de marcas de hipertexto (HTML) con hojas de estilo en cascada (CSS), para estandarizar la estructura de las interfaces generadas, se utilizó el framework Bootstrap en la versión 3, debido a que permite escalar sitios web y aplicaciones de forma fácil y eficiente con una sola base de código, desde teléfonos o tabletas a escritorios con consultas CSS [13].

Para el desarrollo del generador de código se aplicó el método científico experimental, el cual mediante una secuencia de acciones que se llevan a cabo de manera ordenada, permitió establecer las conclusiones sobre el uso

eficiente de las tecnologías actuales para el desarrollo de software ágil. En cuanto al proceso de generación de código se utilizó plantillas de código predefinidas para los tres tipos de escenarios: generación de plantilla de aplicación, generación de modelo y generación de interfaces gráficas que permiten operaciones de: crear, leer, actualizar y eliminar registros (CRUD).

Los proyectos que se desarrollan en Django constan de 4 componentes fundamentales las cuales son: modelo, controlador, enrutador url y vista. En la figura N° 4 se muestra como es la interacción de estas partes.

III. RESULTADOS Y DISCUSIÓN

Con la finalidad de resolver los grandes desafíos a los que se enfrentan los programadores y para facilitar el proceso de codificación, se ha desarrollado un generador de código basado en plantillas. El generador de código permite automatizar tareas de codificación y configuración que son repetitivas, además de resolver los detalles de implementación abstractos, reducir el tiempo de desarrollo y por ende los costos.

Para mostrar las funcionalidades del generador de código se resolvió el desarrollo de una aplicación para el control de inventarios en una empresa, en la figura N° 3 se muestra el diagrama de la base de datos que se va a utilizar para procesar la información

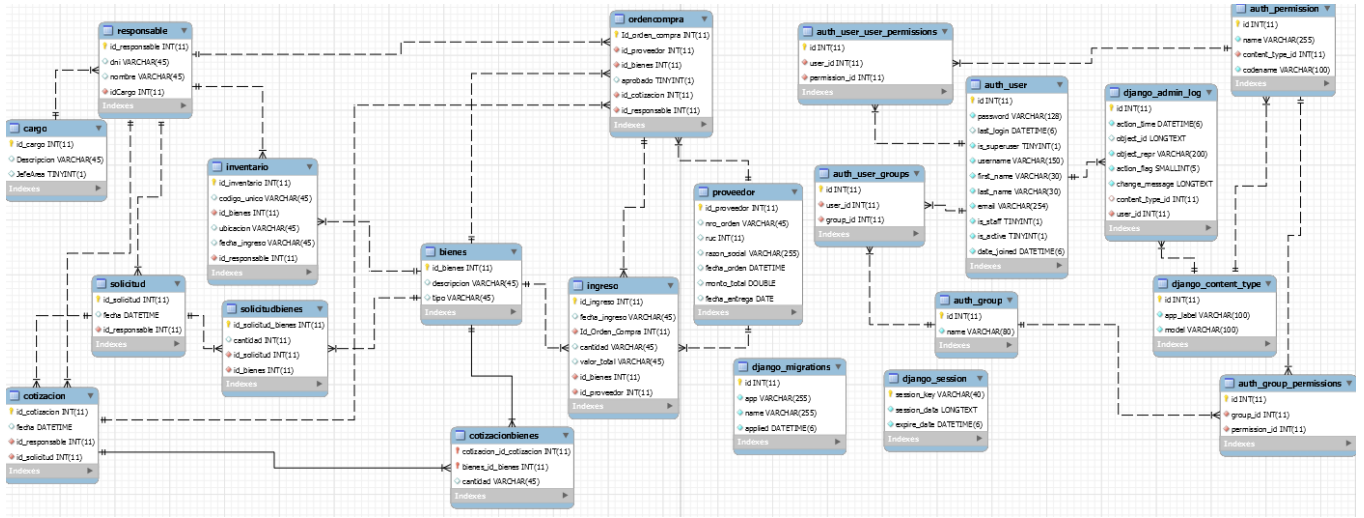


Fig. 3. Diagrama de la base de datos control de inventarios

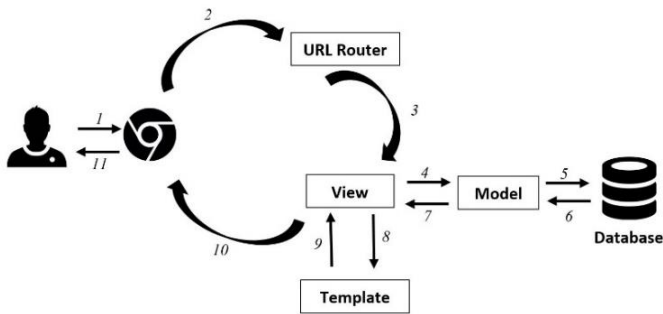


Fig. 4. Componentes de aplicaciones en Django [14]

El proceso de desarrollo de aplicaciones con el framework Django consta de tres partes las cuales son: generación de la aplicación, generación del modelo y generación de las interfaces, que permiten realizar las distintas operaciones CRUD. Con la finalidad de potenciar la usabilidad del generador de código se desarrolló una interfaz gráfica intuitiva que guiará al desarrollador durante el proceso de generación como se puede apreciar en la figura N° 5.

Django code generator App

template, Model and CRUD magical writer

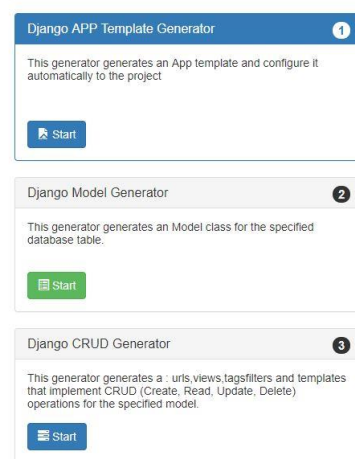


Fig. 5. Interfaz principal del generador de código

Adicionalmente se ofrece la opción de generar código también por línea de comandos como se muestra en la figura N° 6.

```
[djangoappsgenerator]
djangoapps

[djangocrudgenerator]
djangocrud
tests

[sessions]
clearsessions

[staticfiles]
collectstatic
findstatic
runserver

(proyecto_final) C:\django_apps\code_generator>python manage.py
```

Fig. 6. Opciones de generación de código por línea de comandos

En las siguientes secciones se va a detallar las opciones de generación de código que ofrece esta solución.

A. Generación de plantillas de aplicación

El generador de plantillas de aplicación permite crear aplicaciones bases estructuradas, esto es posible mediante la utilización del comando: `djangoapps`, este comando recibe un parámetro, el cual es el nombre de la app. En la figura N° 7, se muestra un ejemplo de la ejecución del comando con su respectivo resultado.

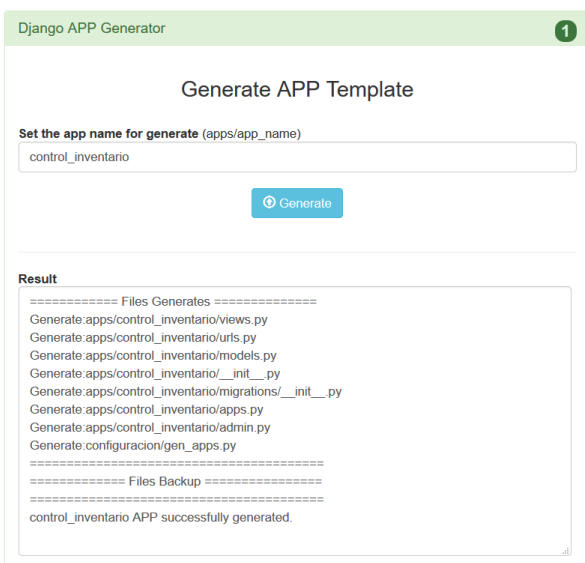


Fig. 7. Ejemplo de generación aplicación plantilla

Las aplicaciones en Django se las debe declarar en el archivo de configuración principal, para que al momento de arrancar el servidor pueda cargar todas las aplicaciones que conforman el proyecto, esta acción la hace el generador declarando la app generada en el archivo `gen_apps.py` de manera automática.

```
1 # -*- coding: utf-8 -*-
2 GEN_APPS = ()
3
4 GEN_APPS += (
5     'apps.control_inventario',
6 )
```

Fig. 8. Declaración de aplicaciones generadas

B. Generación de modelos

Un modelo es la fuente de información sobre los datos [10]. Contiene los campos esenciales y los comportamientos de los datos que está almacenando. Generalmente, cada modelo se asigna a una tabla de base de datos única. El framework Django ofrece el comando `inspectdb`, el cual genera el código fuente de modelo basado en una tabla, la misma que se le debe pasar por parámetro al momento de la ejecución. El generador de código se conecta a la base de datos y muestra las tablas disponibles para seleccionar como se puede apreciar en la figura N° 9.

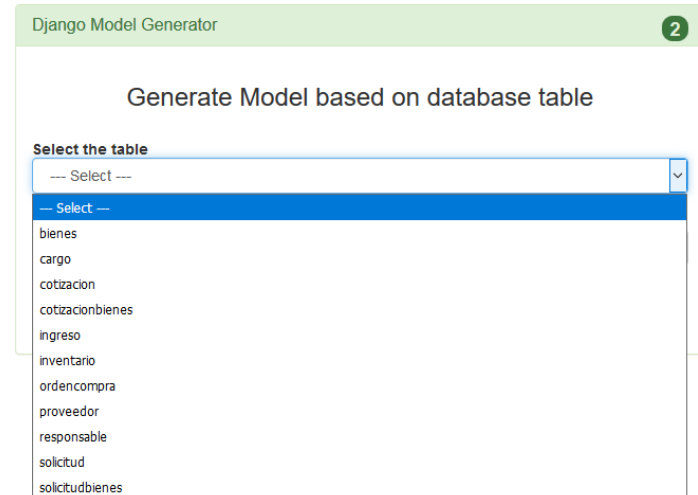


Fig. 9. Tablas disponibles para la generación de modelos

Luego de seleccionar la tabla sobre la cual se quiere generar el modelo se procede a seleccionar la aplicación, el siguiente paso consiste en dar clic en el botón generar para obtener el código fuente.

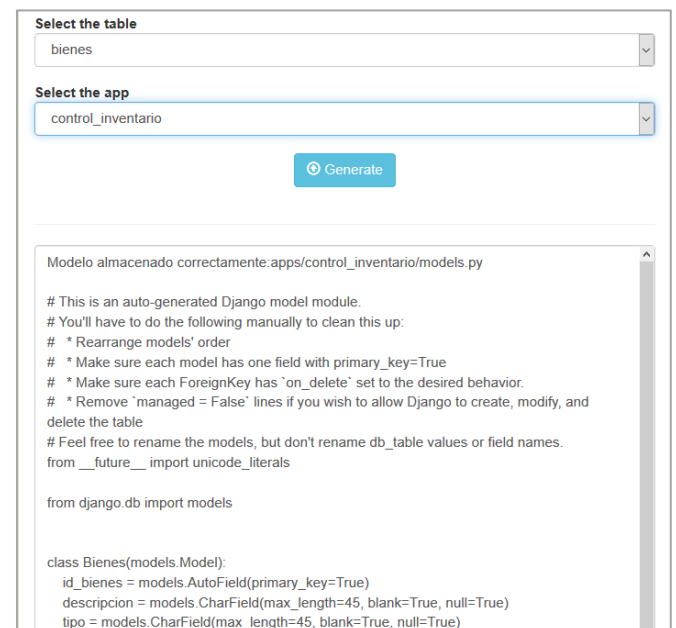


Fig. 10. Generación de modelo bienes

Como se puede apreciar en la figura N° 10, el comando `inspectdb` genera el código fuente del modelo. Para ofrecer la

opción de generar modelos, se utilizó este comando agregando las funcionalidades de almacenar el código fuente en un fichero dentro de la aplicación en conjunto con las funciones: `def __str__(self)` y `def __getitem__(self, item)`. Estas funciones permiten administrar el modelo en las vistas CRUD de una manera más sencilla.

C. Generación de CRUD

Una de las primeras decisiones de diseño de Django fue limitar la funcionalidad del lenguaje de la plantilla. Es muy limitado lo que se puede hacer con las plantillas de Django, que a menudo se piensa que es una cosa muy buena ya que obliga a mantener la lógica de negocio en el lado de Python [15]. En esta fase de generación de código se procedió a generar el controlador (vista) con las funciones CRUD y las respectivas interfaces gráficas con funcionalidad responsiva (plantillas). En la interfaz de generación de código se ofrece la opción de elegir la aplicación y el modelo sobre el cual se quiere generar las operaciones CRUD puesto que una aplicación puede tener varios modelos.

Select the app
control_inventario

Select the model
Bienes

Generate

Result

```

===== Files Generates =====
Generate:apps/control_inventario/templates/control_inventario/bienes_form.html
Generate:apps/control_inventario/templates/control_inventario/bienes_confirm_delete.html
Generate:apps/control_inventario/templates/control_inventario/bienes_list.html
Generate:apps/control_inventario/views.py
Generate:apps/control_inventario/urls.py
Generate:apps/control_inventario/templatetags/control_inventario_filters.py
Generate:apps/control_inventario/templatetags/__init__.py
Generate:apps/control_inventario/__init__.py
Generate:configuracion/gen_urls.py
===== Files Backup =====
Backup:apps/control_inventario/views.py.1497483682
Backup:apps/control_inventario/urls.py.1497483682
=====
control_inventario Bienes CRUD successfully generated.
See the App on: http://localhost:8000/control_inventario/bienes/list/page1
  
```

Fig. 11. Generación CRUD del modelo bienes

En Django, las direcciones o urls deben ser declaradas para que el usuario pueda acceder a las operaciones definidas en el controlador (vista). El generador CRUD al momento de generar las operaciones también genera los enlaces y los almacena en el archivo `urls.py`, este archivo adicionalmente se enlaza al archivo de configuración de urls para aplicaciones generadas. En las siguientes figura N° 12 y figura N° 13, se muestran los dos archivos de configuración de urls.

```

1  #-*- coding: utf-8 -*-
2
3  from django.conf.urls import *
4  from views import *
5
6  bienes_patterns= [
7      url(r'^list/page(?:P<page>[0-9]+)/(\?.*)?$',
8          BienesListView.as_view(),
9          name='list'),
10     url(r'^update/(?:P<pk>\d+)/$',
11         BienesUpdateView.as_view(),
12         name='update'),
13     url(r'^create/$',
14         BienesCreateView.as_view(),
15         name='create'),
16     url(r'^delete/(?:P<pk>[0-9]+)/$',
17         BienesDeleteView.as_view(),
18         name='delete')]
19
20
21 newpatterns = [
22     url(r'^bienes/', include(bienes_patterns,
23         namespace="bienes"))]
24
25
26 try:
27     urlpatterns += newpatterns
28 except NameError:
29     urlpatterns = newpatterns
  
```

Fig. 12. Archivo urls generado por CRUD

```

1  from django.conf.urls import url, include
2  from django.contrib import admin
3
4
5  urls = [
6      # ... your url patterns
7  ]
8
9  urls.append(url(r'^control_inventario/', include('apps.control_inventario.urls', namespace='control_inventario'))
  
```

Fig. 13. Archivo de configuración urls generadas

En la parte inferior del generador CRUD, se muestra la url que el usuario debe visitar para ver el código generado respecto al modelo. En la Fig. 14 se muestra la interfaz principal del código generado para el modelo Bienes en la versión de PC escritorio y teléfonos inteligentes, ya que las interfaces generadas cuentan con la funcionalidad de responsividad.

Bienes

+ New Bienes

Bienes Search 🔍

id To
start end

descripcion

tipo

🔍 Search

id	descripcion	tipo	Actions
1	registro 1	A	Edit Delete
2	registro 2	B	Edit Delete
3	registro 3	B	Edit Delete

Fig. 14. Interfaz Gráfica del modelo Bienes versión PC escritorio

The screenshot shows a mobile application interface for 'Bienes'. At the top, there is a header with 'App' and a menu icon. Below the header, the title 'Bienes' is displayed. A green button labeled '+ New Bienes' is positioned above a search form. The search form includes a search bar and several input fields: 'id_bienes', 'start', 'To', 'end', 'descripcion', and 'tipo'. A blue 'Search' button is located at the bottom of the form. Below the search form, the text 'Number of Bienes: 2' is shown. At the bottom, there is a table with two rows of data. Each row has columns for 'id_bienes', 'descripcion', 'tipo', and 'Actions'. The 'Actions' column contains 'Edit' and 'Delete' buttons for each row.

id_bienes	descripcion	tipo	Actions
1	registro 1	A	Edit Delete
2	registro 2	B	Edit Delete

Fig. 15. Interfaz Gráfica del modelo Bienes versión teléfonos inteligentes

Como se puede apreciar en la Fig. 14 y Fig. 15 el código generado ofrece las funcionalidades CRUD con una interfaz agradable e intuitiva que además es responsiva.

IV. CONCLUSIONES

Con el uso del generador de código se pudo obtener una aplicación web con funcionalidades basadas en buenas prácticas del framework Django, además de interfaces que se adaptan perfectamente a dispositivos inteligentes como teléfonos y tablets. Otro aspecto para tomar en cuenta es que el generador con su interfaz amigable permite crear código fuente de manera inmediata, lo cual significa un ahorro económico al momento de desarrollar proyectos software.

El proyecto de control de inventarios se generó con la utilización del gestor de base de datos MySQL, sin embargo, el generador permite utilizar también otros sistemas gestores de bases de datos como: SQLite o PostgreSQL, puesto que utiliza funciones genéricas del framework lo cual le da un amplio dinamismo.

En trabajos futuros se sugiere agregarle al generador de código la funcionalidad de crear pruebas básicas genéricas que estén basadas en las funciones del controlador (vista), puesto que en esta parte es donde se implementa la lógica del negocio. Adicionalmente sería interesante realizar comparativas con otros frameworks generadores de código para medir el tiempo, calidad y funcionalidad.

REFERENCIAS

- [1] M. Lutz, Learning Python, 5th Edition, O'Reilly Media, 2013.
- [2] L. Pantoja y C. Pardo, «Evaluando la Facilidad de Aprendizaje de Frameworks mvc en el Desarrollo de Aplicaciones Web,» *Publicaciones e Investigación*, pp. 129-142, 2016.
- [3] HotFrameworks, «HotFrameworks,» 22 Julio 2017. [En línea]. Available: <https://hotframeworks.com/languages/python>. [Último acceso: 24 Julio 2017].
- [4] E. M. García, «Desarrollo de una aplicación web para la creación de exámenes de opción múltiple,» 23 Octubre 2013. [En línea]. Available: <https://e-archivo.uc3m.es/handle/10016/18087>. [Último acceso: 10 Julio 2017].
- [5] V. Dmytriiev, «Django Stars,» 2016. [En línea]. Available: <http://djangostars.com/blog/django-or-rails/>.
- [6] K. Shinde y Y. Sun, «Template-Based Code Generation Framework for Data-Driven Software Development,» *In Applied Computing and Information Technology/3rd Intl Conf on Computational Science/Intelligence and Applied Informatics/1st Intl Conf on Big Data, Cloud Computing, Data Science & Engineering (ACIT-CSII-BCD)*, pp. 55-60, 2016.
- [7] RogueWave Software, «University of Cambridge Study: Failure to Adopt Reverse Debugging Costs Global Economy \$41 Billion Annually,» 1 Septiembre 2013. [En línea]. Available: <https://www.roguewave.com/company/news/2013/university-of-cambridge-reverse-debugging-study>. [Último acceso: 2 Julio 2017].
- [8] M. Kanat-Alexander, Code Simplicity: The Fundamentals of Software, O'Reilly Media, Inc, 2012.
- [9] S. Cass, «The 2016 Top Programming Languages,» 26 Julio 2016. [En línea]. Available: <http://spectrum.ieee.org/computing/software/the-2016-top-programming-languages>.
- [10] Django Software Foundation, Django Documentation, 2017.
- [11] L. Casillas, M. Gibert y O. Pérez, Bases de datos en MySQL, Universitat Oberta de Catalunya, 2014.
- [12] E. Arab y A. Díaz, «Impacto de las redes sociales e internet en la adolescencia: aspectos positivos y negativos,» *Revista Médica Clínica Las Condes*, vol. 26, n° 1, pp. 7-13, 2015.
- [13] Bootstrap, «Bootstrap,» 2017. [En línea]. Available: <http://getbootstrap.com>.
- [14] V. Dmytriiev, «Django or Rails,» 2016. [En línea]. Available: <http://djangostars.com/blog/django-or-rails/>.
- [15] D. Roy y A. Roy, Two Scoops of Django: Best Practices for Django 1.8, 2015.
- [16] A. Oliveros, F. Danyans y M. Mastropietro, «Prácticas de Ingeniería de Requerimientos en el desarrollo de aplicaciones Web,» *Facultad de Ingeniería y Ciencias Exactas, UADE*, 2014.
- [17] B. Boehm, C. Abts, W. Brown, S. Chulani, B. Clark, E. Horowitz, R. Madachy, D. Reifer y B. Steece, Software Cost Estimation with Cocomo II, Financial Times Prentice Hall, 2009.

Análisis de sentimiento con Lingmotif: Aproximación metodológica al análisis de intervenciones en series temporales de la Valencia*

Fernández-Cruz, Javier^{1,3}, Baixauli-Pérez, Cristóbal²

javier.fernandez@pucese.edu.ec, cristobal.baixauli@pucese.edu.ec

^{1,2}Pontificia Universidad Católica del Ecuador Sede Esmeraldas – Esmeraldas 080101 Ecuador;

¹Escuela de Lingüística Aplicada, ²Escuela de Ingeniería de Sistemas y Computación

³Universidad de Málaga, Facultad de Filosofía y Letras – Málaga 29071 España

Resumen – Este trabajo pretende describir una aproximación metodológica al desarrollo de análisis de intervención, aplicados a la evolución de la valencia del sentimiento en la Web. Es decir, asentar las bases teóricas necesarias para plantear futuros análisis de series temporales, gracias a los cuales sea posible establecer relaciones causa-efecto entre los cambios estructurales en la evolución de la opinión pública y los posibles factores externos influyentes. Inicialmente se describen los conceptos fundamentales del Análisis de Sentimiento de textos digitales, también conocido como Minería de Opinión. Se describen diferentes técnicas de extracción de datos, así como conceptos ligados al sentimiento como la polaridad y la valencia de los textos. Se indican, a su vez, diferentes técnicas usadas en el análisis de sentimiento como la basada en lexicón, el aprendizaje automático o los métodos mixtos. Por otro lado, se describe en detalle la metodología desarrollada en el software Lingmotif, gracias al cual podemos obtener diferentes métricas de sentimiento de forma sencilla. Posteriormente se describen teóricamente las series temporales y especialmente sus cambios estructurales. Diferenciando entre cambios endógenos de estacionalidad y cambios estructurales propios de intervenciones exógenas a la propia dinámica de la opinión pública.

Palabras clave: Análisis de Sentimiento, Minería de Opinión, Lexicón, Series Temporales, Cambios Estructurales, Análisis de Intervención.

I. INTRODUCCIÓN

¿Ha influido la actual crisis en la forma en la que nos sentimos en relación con el lenguaje? En la última década se han producido profundos cambios en el mundo de la economía y las finanzas, así como en las subjetividades de la población general. Estos cambios quedan reflejados en el discurso de la prensa, los poderes económicos y la ciudadanía. El lenguaje es un ente vivo que refleja esta evolución constante y arrastra la actitud de los sujetos a través del lenguaje evaluativo: ideas, opiniones y creencias situadas en el núcleo de la subjetividad en el lenguaje humano [1]. Las distintas herramientas de procesamiento del lenguaje natural (PLN, o en inglés, NLP) han resultado ser muy útiles para observar la evolución de determinadas opiniones (el desempeño de un artista, los sucesos políticos, etc.) a lo largo de un intervalo de tiempo y conocer sus variaciones. Este trabajo propone un modelo para el estudio a partir de líneas temporales para analizar la evolución de las

valencias de una selección de léxico relacionado con la crisis.¹

La crisis financiera mundial estalló en 2007 causando la peor situación económica desde la Gran Depresión provocando el colapso del sistema financiero global. Con la quiebra de Lehman Brothers en 2008, el auge de la crisis de la deuda soberana en Europa, o el nuevo capítulo de la situación política en América Latina, ha puesto a las clases no privilegiadas en una situación de sufrimiento generalizado. A partir de esta línea de sucesos, la prensa se ha hecho eco de los cambios lingüísticos que han surgido a partir de su debate, y han aparecido expresiones nuevas o ha variado la valoración que tenemos con respecto a determinados términos. Para ilustrar esto, el término “squeezed middle” (la clase media exprimida) fue término del año para el diccionario Oxford en 2011 y el uso de la palabra “deuda” se convirtió en uno de los principales mantras de la campaña presidencial en el Ecuador en 2017. Por ello, podemos esbozar que el ascenso de terminología de dominios lingüísticos muy específicos como el económico/financiero hacia el lexicón del hablante no especializado está íntimamente vinculado a los efectos directos de la pérdida de calidad de vida por la crisis [2, 3].

La combinación del desarrollo tecnológico que en la presente década ha duplicado la cantidad de información publicada a diario, por un lado, y la radical transformación sociocultural que esto ha supuesto, por otro, genera nuevos desafíos a la hora de estudiar la influencia de la actualidad en el lenguaje y viceversa. La profusión que ha fomentado la Web 2.0 convierte a Internet en una suerte de ágora global: el discurso de los usuarios tiene hoy mejores posibilidades de ser visibilizado por sus pares y de generar tendencias. De esta manera, los usuarios destacados pueden disputar en términos de influencia con las cabeceras editoriales. Ya sea la expresión de los distintos argumentarios políticos, la recomendación de hábitos y preferencias de consumo o el compartir conocimientos y experiencias, es fácil extraer estos datos y analizar estas opiniones. Gracias al análisis textual se pueden medir las distintas tendencias y examinar múltiples factores: desde la evolución de la opinión pública a la evolución del uso de distintas palabras o expresiones.

¹ Esta investigación se ha llevado a cabo gracias a la financiación aportada por el Concurso de Investigadores de la PUCE Esmeraldas 2016-2017

Las instituciones públicas o privadas tienen la necesidad de observar las tendencias, con vistas a conocer el sentimiento de su audiencia objetivo (ciudadanos, clientes, etc.) y su relación con los servicios o productos que ofrecen. Tradicionalmente, este análisis se ha venido realizando por parte de analistas de mercado, medios y relaciones públicas mediante sesudos análisis “manuales” de la prensa, informes realizados ad hoc, análisis de mercado, reseñas de clientes y encuestas de satisfacción. Estos enfoques tradicionales son complementados y en muchas ocasiones superados por la inmediatez que ofrecen redes sociales o las cabeceras digitales. Gracias a este enfoque, podemos obtener datos a partir de fuentes web mediante el uso de arañas (o, en inglés, crawlers) o las interfaces de programación de aplicaciones (siglas en inglés, API) disponibles en la mayor parte de webs sociales como Twitter, que proveen a los usuarios y desarrolladores de arquitecturas abiertas que permiten la extracción automatizada de contenidos. Las arañas y las APIs permiten obtener en minutos grandes cantidades de datos que incluyen muestras idiomáticas procedentes de cualquier tipo de dominio. Los textos extraídos a partir de la web y sus herramientas de análisis automatizado han sido el objeto de una profunda labor de investigación durante la última década y nuevas disciplinas han surgido a partir de su desarrollo. Una de las más prominentes es el análisis textual, del cual el análisis de sentimiento es quizá el subcampo más relevante dentro de la lingüística aplicada. A través de distintos algoritmos, podemos calcular la orientación semántica de las palabras, expresiones multipalabra y oraciones.

La recopilación y almacenamiento de gran cantidad de datos gracias al uso de arañas o APIs hace posible el análisis del sentimiento de determinados elementos léxicos a lo largo del tiempo. El análisis de series temporales se muestra muy útil para interpretar los fenómenos asociados a la dinámica del sentimiento en las redes mediante la captura total o parcial de las características relevantes del fenómeno subyacente. En la literatura se han propuesto gran variedad de modelos, p.e. [4, 5, 6, 7]. La utilidad o no de estos para el caso del análisis del sentimiento radica en la capacidad que tengan para reproducir la dinámica del proceso generador de la serie. Tomando las precauciones necesarias para controlar los diferentes cambios estructurales que pueda tomar su dinámica interna, como cambios en su estacionalidad. Los cuales deben ser identificadas para definir eficazmente los modelos autorregresivos y entender con precisión la evolución de la valencia u otras métricas de sentimiento.

Por otro lado, y de mayor interés, son la identificación y análisis de sus cambios estructurales producidos por intervenciones exógenas. Un cambio estructural en una serie se presenta cuando hay modificaciones inesperadas permanentes o instantáneas. El análisis de intervención es la aplicación de procesos que incluyen los efectos de fuerzas exógenas o intervenciones en el análisis de la dinámica temporal. En otras palabras, estas intervenciones obedecen a incidentes de distinta naturaleza, por ejemplo: cambios políticos, cambios en los precios, períodos de campaña

electoral, agenda setting, campañas publicitarias, desastres naturales, atentados o guerras, entre otros. Pudiendo usar estas técnicas para identificar aquellas fuerzas exógenas que tengan una relación causa-efecto demostrada estadísticamente.

Este trabajo se lleva a cabo en el contexto del proyecto LexiEcon desarrollado en la PUCE Esmeraldas que busca expandir trabajos anteriores [8] y otros [10, 11, 12, 13, 14, 15, 16, 17, 18, 19] realizados por el grupo de investigación Tecnolengua de la Universidad de Málaga.

II. DESARROLLO

A. *Análisis de Sentimiento*

El análisis de sentimiento (*Sentiment Analysis*), a veces referido como minería de opinión (*Opinion Mining*), es una disciplina que surge desde el procesamiento del lenguaje natural y la lingüística computacional y se dedica al tratamiento automatizado de la opinión y la subjetividad de los textos. El desarrollo de esta disciplina permite la clasificación del sentimiento en textos completos o secciones ya sean de tipo positivo o negativo. Sus aplicaciones reciben una entrada de texto electrónico de cualquier tipo y estas devuelven un resultado de sentimiento general con mayor o menor sensibilidad y, frecuentemente, el análisis textual correspondiente. Los distintos desarrollos en el campo del análisis de sentimiento se aplicado con éxito en múltiples campos: profesional, comercial o académico, etc.

La fundación del análisis del sentimiento como disciplina se remonta a una fecha próxima al inicio del siglo XXI, pero la investigación relacionada con la opinión y el lenguaje evaluativo puede trazarse a la década de 1970 y distintas teorías lingüísticas han estudiado el lenguaje evaluativo asignándole distintos nombres: connotación [20], afecto [21], actitud [22], postura [23] o valoración [24], entre otros [25, 26]. Por otro lado, especialmente durante la presente década, se han creado un sinfín de estudios de este tipo con enfoque en inteligencia artificial, siendo uno de los primeros el modelo de comprensión subjetiva que ya a finales de la década de 1970 se desarrolló en el MIT para comenzar a estudiar el discurso y la ideología de manera automatizada [27].

En relación con el estudio de la opinión, toda la información que se encuentra en un texto de cualquier categoría se puede dividir en información factual y opiniones. Se entiende como información factual a las expresiones de carácter objetivo que describen a cualquier tipo de entidad junto con sus subatributos. Mientras tanto, las opiniones incluyen las expresiones subjetivas que describen a cualquier tipo de emoción en un sujeto determinado, sus sentimientos y las evaluaciones que realice sobre cualquier entidad [28]. En mayor medida, las opiniones expresan siempre alguna entidad que suele contener distintos subcomponentes jerárquicos. La subdivisión de mayor orden es el objeto, el cual se define como cualquier entidad en el mundo real (p.e. una persona, un producto, una idea...). Existen patrones de sentimiento más complejos [29, 30]. Para nuestro propósito, este modelo puede simplificarse en un esquema de partes y

características [31, 32]. Las partes se encargan de las distintas subdivisiones físicas o abstractas de un objeto determinado, mientras que las características se encargan de sus variables. Para ilustrar esto con un ejemplo sencillo, tomemos como objeto a una naranja, la cual tiene una serie de subdivisiones que contienen sentimiento. Las distintas partes de una naranja serían p.e. cáscara, pulpa, jugo y semillas. Entre las características de la naranja incluiríamos color, aroma, sabor, tamaño, etc. Por ejemplo, en un marco de opinión relacionado con los gustos del consumidor estándar, una naranja descrita con ‘semillas grandes’ tendrá una orientación semántica negativa si la comparamos con una naranja con ‘semillas pequeñas’.

B. Clasificando la orientación semántica: polaridad vs valencia²

La orientación semántica relacionada con una determinada opinión sobre una característica determinada muestra si la opinión es positiva, negativa o neutral. Los sistemas de análisis de sentimiento usan distintos métodos para asignar los valores de orientación semántica, sean estos positivos o negativos, a la opinión de las palabras o unidades multipalabras [28, 29, 33]. Dos son los enfoques principales utilizados a la hora de clasificar textos según su sentimiento. El primero, es el enfoque Thumbs Up/Thumbs Down (pulgares arriba/pulgares abajo) [34, 35] el cual asigna valores de polaridad a las unidades de sentimiento a través de una clasificación binaria simple. Por ejemplo, a la palabra ‘excelso’ se le asignará un simple valor “+” o “POSITIVO” y a la palabra ‘vomitivo’ se le asignará un valor “-” o “NEGATIVO”. Por otro lado, el segundo enfoque, Seeing stars (viendo estrellas) [36] tiene mayor granularidad y está basado en la forma en la que los usuarios asignan calificaciones a los productos en páginas de comercio electrónico como es el caso de p.e. Mercadolibre o Amazon, gradando el sentimiento de las palabras o expresiones multipalabras mediante la asignación de grados o puntos en una escala preestablecida. En este caso, a la palabra ‘excelso’ se le asignará p.e. una puntuación 10/10, mientras que la ‘palabra “aceptable” recibiría una puntuación de “7/10” en dicha escala decimal. Por otro lado, a la palabra ‘catástrofe’ se le asignará una puntuación de “1/10” [36]. Además de valores positivos o negativos, muchos sistemas de análisis de sentimiento toman en cuenta la neutralidad y unidades multipalabras como “bolsita de té” simplemente tendrán una asignación de valor “neutral” (equivaldría a una puntuación de 5/10 en el modelo Seeing Stars). La precisión con la que se desgrane un sistema de análisis de sentimiento es una variable importante ya que sistemas como Lingmotif consideran la presencia de segmentos neutrales para realizar el cálculo pormenorizado del sentimiento de un texto.

C. Modificadores contextuales de valencia

La orientación semántica de muchos de los términos en un texto no suele cambiar de valencia. Los modificadores

contextuales de valencia (Contextual Valence Shifters o CVS) se describieron originalmente en [37], y fueron implementados en inglés por [38] en su CLaC System, y por [39] en la Semantic Orientation CALculator (SO-CAL). A partir de este sistema, la valencia de otras palabras con carga sentimental cambia dependiendo de las palabras que le rodeen. Hay un conjunto de fenómenos léxicos que causan la inversión de la valencia de un término puntual. Puede haber un cambio de orientación semántica, o quedar este neutralizado por completo. En términos simples, ocurren aquí dos fenómenos distintos: la inversión de la valencia o movimientos en la puntuación dentro de la escala de graduación. Cuando nos referimos a la inversión, la polaridad de uno de los términos cambia de una polaridad determinada a su opuesto o se neutraliza debido a la presencia de un término en su proximidad. Los casos más

Tabla 1. Ejemplo de asignación de polaridad/Valencia según los enfoques Thumbs up / Thumbs Down y Seeing Stars

	<i>Enfoque Thumbs up – Thumbs down</i>	<i>Enfoque Seeing stars</i>
CATÁSTROFE	NEGATIVO	1/10
TRASTADA	NEGATIVO	4/10
ACEPTABLE	POSITIVO	7/10
EXCELSO	POSITIVO	10/10

típicos son las negaciones, como podemos ver en el siguiente ejemplo, un término positivo (‘inteligente’) invierte su polaridad negativa debido a la influencia de la palabra ‘no’ que le premodifica.

no(+SER) + sust. neg.
“No es estúpido”

El segundo fenómeno se encarga del cambio progresivo de valencia debido a la proximidad de otro elemento léxico en su contexto. Dependiendo del caso, estos elementos léxicos pueden intensificarse o suavizarse y en consecuencia cambiará la valencia de una palabra con carga sentimental. Por ejemplo, usar “del todo” como complemento del adjetivo negativo ‘tonto’ en “No es tonto del todo” suaviza la valencia negativa inicial de dicho adjetivo [15].

D. Aplicaciones de análisis del sentimiento

Para dar crédito a algunas de las soluciones de análisis de sentimiento más destacables, ya sean de tipo comercial o académicas, nos gustaría destacar a SentiWordNet [40], un

² Aunque los términos polaridad y valencia se utilizan de manera casi intercambiable en mucha de la literatura cuando nos referimos al cálculo del sentimiento en los textos evaluativos, Moreno-Ortiz y Pérez-Hernández

(2013) hacen diferencia en su uso. Polaridad en este texto se usa de manera binaria (positivo ó negativo), mientras que valencia se usa para términos en escala (1-10/10).

lexicón de polaridad de código abierto que ha servido como base de una gran cantidad de sistemas posteriores [41, 42, 43]. Por otro lado, aplicaciones de análisis de sentimiento que se escapan del dominio de la investigación académica y se distribuyen de manera comercial. Como, por ejemplo, Lexalytics. Esta aplicación web está basada en aprendizaje automático y toma como base un lexicón de palabras individuales y expresiones multipalabras anotadas. Según indica en su web, la aplicación cubre 21 lenguas distintas y devuelve una puntuación final similar al modelo seeing stars para las distintas entidades y para texto completo. Además, presenta una API personalizable, detecta entidades y temáticas en el texto y maneja terminología de distintos dominios [44].

E. Análisis de series temporales con cambios estructurales

El estudio de la dinámica de la opinión pública más allá de la utilidad clara para multitud de empresas o gobiernos está ganando relevancia en muchos ámbitos en los últimos años. Cada vez son más los artículos que se publican sobre temas relacionados. La modelización de la dinámica de la opinión pública mediante series temporales puede realizarse gracias a la agregación de la puntuación de sentimiento de texto (TSS, Total Sentiment Score) obtenida para un conjunto preestablecido de fuentes de datos. La naturaleza y cantidad necesaria de estos textos dependerá en cada caso de las dinámicas concretas que quieran ser abordadas.

La diferencia entre identificar correctamente un modelo o no, radica principalmente en identificar los diferentes fenómenos concurrentes en la serie de tiempo. A si como identificar cada uno de sus cambios estructurales y su naturaleza.

Un cambio estructural perturba la componente determinista de la serie temporal, provocando una pérdida de poder en la identificación de la dinámica subyacente. Tal como indica Balke [45], los cambios afectan a la estructura de autocorrelación, por tanto, también afectan a la estimación de la función de autocorrelación (ACF), al autocorrelograma parcial (PACF), y las funciones autoextendidas de autocorrelación (EACF), dificultando la identificación del modelo y dando como resultado parámetros inconsistentes.

Perron [46] analiza los cambios estructurales como endógenos, planteando que los eventos aleatorios pueden tener un efecto permanente en la serie cuando estas tienen raíz unitaria. Las propiedades de la dinámica y las técnicas aplicables a cambios estructurales han sido también consideradas por J. Bai [47]. Por otro lado, Atkinson et al. [48] desarrollan técnicas para identificar cambios de nivel y la estimación de parámetros ARIMA (Autoregressive Integrated Moving Average). Harvey et al. [49] consideran pruebas para detectar outliers aditivos e innovacionales ante cambios estacionales en la media. La mayor parte de la literatura posterior ha seguido centrándose en el modelado de cambios estructurales en series lineales, dando poca importancia al caso no lineal.

A continuación, pasamos a describir dos tipos principales de cambios estructurales; por un lado, los cambios estructurales de estacionalidad, aquellos que afectan a los valores esperados de la serie en una pauta periódica. Y, por otro lado,

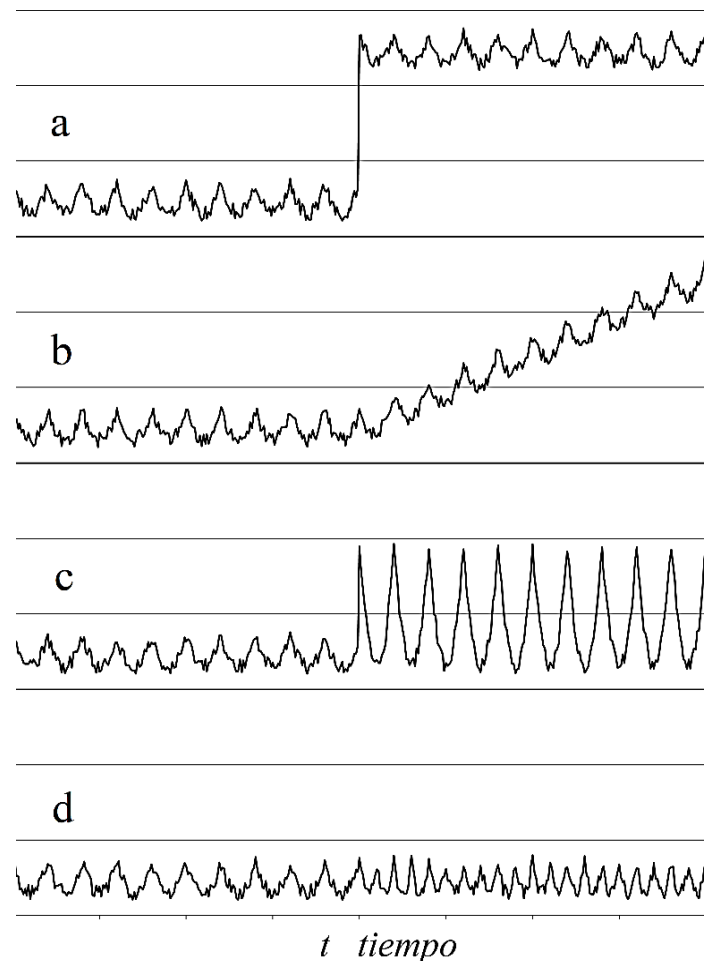


Fig 1. Cambios estructurales de estacionalidad: a) nivel; b) tendencia; c) amplitud y d) frecuencia

los cambios estructurales por intervenciones, aquellos que afectan a la serie como consecuencia de una fuerza exógena.

F. Cambios estructurales de estacionalidad

La estacionalidad en una serie de tiempo es aquella característica que hace que determinados patrones se repitan con una periodicidad determinada. Según Kaiser y Maravall [50] un cambio simple en la componente estacional de la serie puede afectarla modificando su nivel, tendencia, amplitud o frecuencia tal como se presenta en la figura 1. Pudiendo darse cambios múltiples de diferente naturaleza.

G. Cambios estructurales por intervención

El análisis de Intervención para series temporales Y_t modela los cambios estructurales a través de una componente aditiva que representa la intervención exógena, la dinámica de la serie se describe del siguiente modo:

$$Y_t = f(k, X_t, t) + Z_t \quad (1)$$

Donde la componente regular de la serie viene representada por un ruido estocástico Z_t , equivalente al de los modelos ARMA donde

$$Z_t = \pi(B) = \frac{\theta(B)}{\varphi(B)\alpha(B)}$$

Y el efecto de las irregularidades viene representado por la función $f(k, X_t, t)$ que depende del tiempo t , y de los efectos de las intervenciones (variables exógenas) X_t , y de un conjunto de parámetros desconocidos k . Donde las variables exógenas siguen la siguiente dinámica:

$$f(\delta, \omega, X_t, t) = \sum_{j=1}^k \frac{\omega(B)}{\delta(B)} X_t^{(j)}$$

donde los parámetros desconocidos k son sustituidos por ω y δ . Esta dinámica puede ser simplificada por el modelo de función general de transferencia dado por

$$Y_t = Z_t + \frac{\omega(B)}{\delta(B)} X_t^{(i)} \quad (2)$$

donde $Z_t = \pi(B)$ es el ruido estocástico y $X_t^{(i)} = \{1, t = i; 0, \text{de lo contrario}\}$ es un booleano que indica la ocurrencia o no de la intervención.

La dinámica de la opinión pública, así como el estudio de cualquier dinámica en la naturaleza, está potencialmente influenciada por una gran variedad de fuerzas exógenas, las cuales provocarán cambios estructurales atípicos de diferente tipo, magnitud, posición y duración.

Los principales modelos atípicos usados en el análisis de intervención tal como muestra la figura 2 son:

AO, el valor atípico aditivo (en inglés additive outlier) se refiere a un valor transitorio con forma de pulso situado en la t -ésima observación tal como muestra;

LS, el cambio estructural de nivel (level shift), es un abrupto pero permanente escalón o salto, que genera un cambio de ω en la serie causado por una intervención;

TC, el cambio estructural temporal (temporary change) es una intervención que tiene un efecto ω en el tiempo t_1 , que se desvanece gradualmente y

IO, el cambio estructural innovacional (innovational outlier) es un cambio inesperado en un tiempo t que influye en Y_t, Y_{t+1}, \dots

La definición formal de estos cuatro modelos parte de la ecuación general

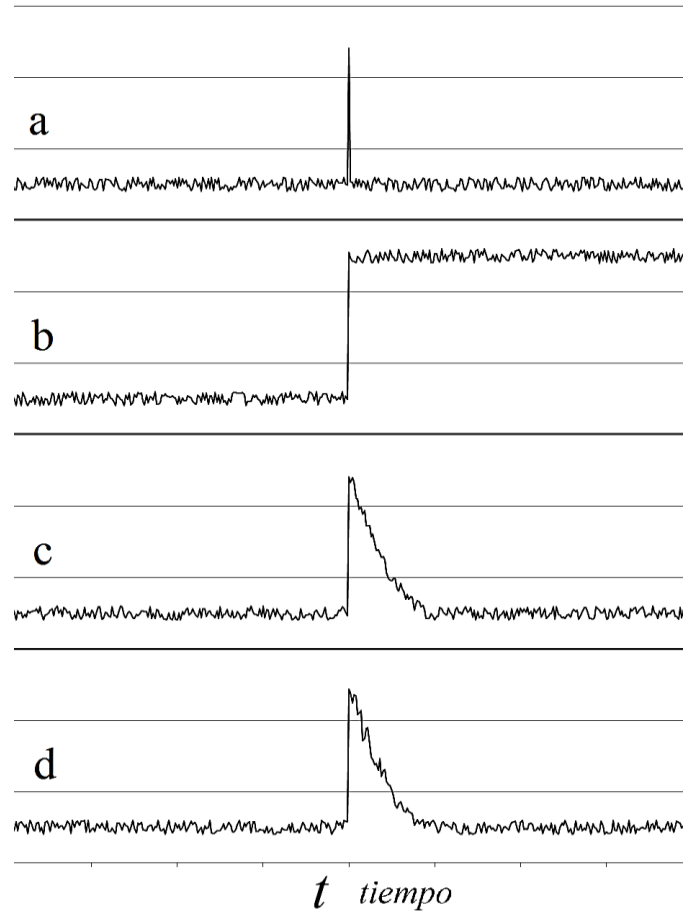


Fig 2. Cambios estructurales por intervención: a) AO Additive Outlier; b) LS, Level Shift; c) TC, Temporary Change; d) IO, Innovational Outlier

$$Y_t = Z_t + \frac{\omega}{1-\delta B} X_t^{(i)} \quad (3)$$

siendo el parámetro δ el que finalmente caracteriza la dinámica particular de cada modelo tal como se detalla en la tabla 2. Y cerrando la definición de los mismos con su dinámica residual.

Tabla 2. Modelos de Intervención para cambios estructurales exógenos: AO, valor atípico aditivo; LS, cambio de nivel; TC, cambio temporal; IO, cambio estructural innovacional. Dinámica y residuo de los modelos.

	Parámetro	Dinámica Y_t	Residuo e_t
AO	$\delta = 0$	$Z_t + \omega X_t^{(i)}$ (4)	$\omega\pi(B)X_t^{(i)} + a_t$
LS	$\delta = 1$	$Z_t + \frac{\omega}{1-B}X_t^{(i)}$ (5)	$\omega \frac{\pi(B)}{1-B}X_t^{(i)} + a_t$
TC	$0 < \delta < 1$	$Z_t + \frac{\omega}{1-\delta B}X_t^{(i)}$	$\omega \frac{\pi(B)}{1-\delta B}X_t^{(i)} + a_t$
IO	-	$\pi(B)(\omega X_t^{(i)} + a_t)$ (6)	$\omega_0 X_t^{(i)} + a_t$

El Análisis de Intervención ayudará a identificar las diferentes variables exógenas responsables de los cambios estructurales. Extrayendo la dinámica aportada por estos cambios estructurales, se asume que la dinámica restante puede explicarse mediante un modelo autorregresivo de media móvil ARMA, pudiendo estimar sus parámetros gracias a procedimientos de mínimos cuadrados o estimaciones de máxima verosimilitud.

III.DISEÑO

Los distintos enfoques del análisis de sentimiento pueden resumirse de manera rápida en tres: un enfoque basado en lexicón, un enfoque basado en aprendizaje automático y sus híbridos. Los sistemas basados en aprendizaje automático hacen uso de corpus como datos de entrenamiento para así clasificar un texto de entrada como positivo o negativo. Por otro lado, los enfoques basados en lexicón dependen de diccionarios de emociones donde a las distintas entradas léxicas se les ha asignado una polaridad o valencia determinada.

Mientras que este enfoque léxico no requiere de datos etiquetados para ofrecer resultados precisos, depende excesivamente de poderosas fuentes textuales que no siempre están disponibles y que con frecuencia dependen del dominio en el que esté especializado el vocabulario del texto. Por el contrario, los sistemas basados en aprendizaje automatizado no requieren diccionarios. Aunque ofrecen buenos resultados de precisión en los dominios especializados, los niveles de precisión son más bajos en textos con lenguaje generalista. Entonces, para realizar esta tarea, se requiere una selección representativa de textos de entrenamiento etiquetados previamente acompañados de un conjunto de parámetros de tipo selectivo [51].

Con independencia de la técnica usada para dotar los textos de métricas de sentimiento, los requerimientos estadísticos de los datos que alimenten el análisis temporal dependerán de la volatilidad temporal de los fenómenos que se quieran identificar. Así como de la procedencia de los textos, tweets, noticias... y el impacto que hayan tenido estos en las redes.

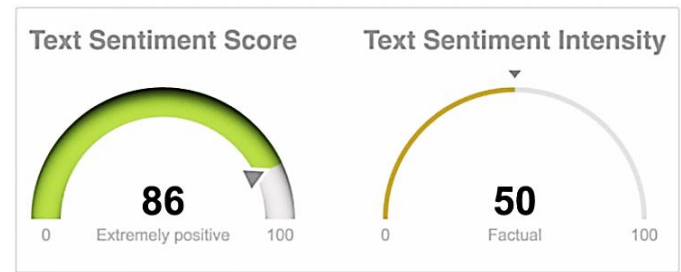


Fig 3. Pantalla de marcadores de puntuación de Sentimiento de Texto (TSS) e Intensidad de Sentimiento (TSI)

El equipo Tecnolengua comenzó a partir del desarrollo de Sentitext, un sistema de análisis de sentimiento para la lengua española [10,11,52] y evolucionó con el proyecto Lingmotif para integrar además el análisis del inglés. Ambos comparten principios de análisis: una cimentación firme sobre un conjunto de recursos léxicos como alternativa de peso a un complejo conjunto de algoritmos. En las figuras 3 y 4 puede verse respectivamente la pantalla de marcadores de puntuación de Sentimiento de Texto (TSS) e Intensidad de Sentimiento (TSI) y la interfaz de Lingmotif 1.0 para Windows.

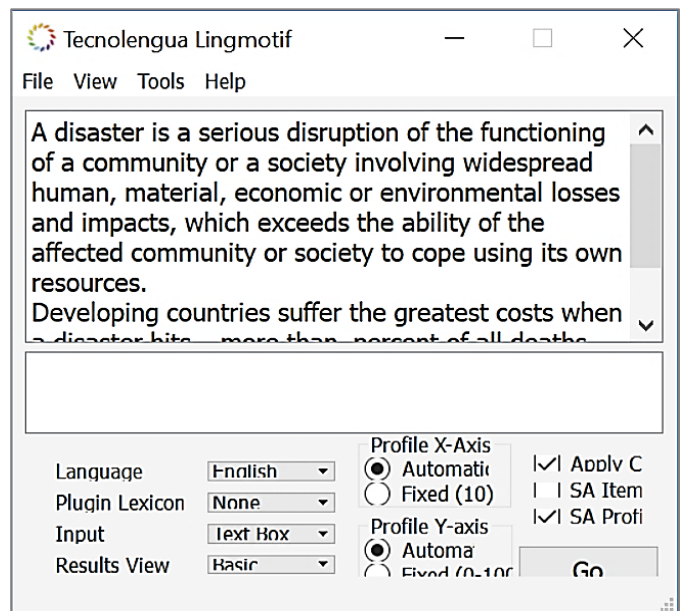


Fig 4. Interfaz de Lingmotif 1.0 para Windows

El proceso de análisis se abarca un conjunto de nueve pasos [18], y resumimos los principales aquí:

- 1.El texto de entrada es preprocesado, tokenizado, lematizado, y etiquetado gramaticalmente. Asimismo, se detectan y etiquetan las expresiones multipalabras.
- 2.Se buscan las palabras con significado léxico y las expresiones multipalabras en los lexicones de sentimiento. Si se encuentran, se les asigna la valencia correspondiente.
- 3.Se realiza una búsqueda de reglas contextuales para cada una de las palabras léxicas o expresión multipalabras. Se les

asigna una valencia resultante a los segmentos con coincidencias en la búsqueda a partir de la aplicación de las reglas contextuales.

4. Se calcula la intensidad de sentimiento del texto (*Text Sentiment Intensity*, la proporción de unidades con carga sentimental vs. unidades neutrales) que devuelve un valor en el rango de 0-100.

5. Se calcula la puntuación de sentimiento del texto (*Text Sentiment Score*) (*Global Sentiment Value*, o TSS) que devuelve un valor en el rango de 0-100.

Los buenos resultados de rendimiento de esta herramienta dependen de un lexicón con alto nivel de detalle, alta calidad y amplia cobertura. Lingmotif hace uso de tres fuentes de datos lingüísticos: un diccionario de palabras individuales (13.250 entradas en inglés), un segundo diccionario de expresiones multipalabras (12.300 palabras en inglés), y un amplio juego de 750 reglas contextuales, una implementación de los modificadores contextuales de valencia. El análisis se lleva a cabo mediante la identificación de las palabras y frases que están incluidas en el lexicón. La puntuación total del texto se computa como una función de las puntuaciones positivas, negativas y neutrales acumuladas. Existe la posibilidad de aplicar lexicones de un dominio lingüístico específico a partir de archivos CSV que actualice las valencias del lexicón principal. En nuestro caso, el equipo LexiEcón de la PUCE Esmeraldas está desarrollando un lexicón del dominio económico y financiero que, en la versión actual, incluye 3.100 palabras y 3.000 expresiones multipalabras en lengua inglesa, obtenidas a partir de a través de un corpus que recopila una selección de noticias económicas entre 2007 y 2015.

Esto se implementa en una aplicación de escritorio multiplataforma (Windows, MacOS y GNU/Linux) desarrollada en Python y con una interfaz amigable para el usuario. Ofrece información visual del perfil de sentimiento de los textos introducidos a partir de gráficos, lo que permite que los usuarios puedan comparar directamente el perfil de una gran cantidad de archivos, y además, puede procesar una serie de documentos en un orden determinado. Estas características facilitan la tarea del analista a la hora de determinar los cambios de sentimiento en uno o varios textos, así como observar la evolución del sentimiento a partir de múltiples archivos [15]. Los resultados se visualizan en un archivo HTML local que muestra de manera amigable con el usuario la puntuación, la gráfica y un análisis detallado del texto. Además, estos datos se almacenan en un documento XML descargable que contiene toda la información relevante para ser procesado con facilidad en cualquier otra aplicación.

IV. CONCLUSIONES

El cálculo automático de la orientación semántica puede aportar una panorámica privilegiada de la opinión pública sobre ciertos asuntos – reseñas de consumidores, niveles de aceptación política, etc. Sin embargo, los niveles de precisión dependen en muchas ocasiones de amplios volúmenes de datos. El lema utilizado por Sinclair más grande siempre es mejor también hace alusión al análisis de sentimiento:

mientras más datos tenga un sistema, mayor probabilidad de resultados de sentimiento acertados. Sin embargo, un análisis cualitativo manual de los resultados puede detectar excepciones y fallos de detección no observados por el sistema automático. Los problemas más frecuentes descritos en la literatura [1, 33] en los sistemas de análisis de sentimiento son los siguientes:

1. Dominio: una palabra con sentimiento positivo o negativo puede tener una orientación opuesta en un contexto especializado diferente. “Mucha mierda para tu actuación de esta noche” en un dominio general considerado como una expresión negativa, sin embargo, en el dominio de las artes escénicas es considerado un enunciado con una orientación altamente positiva.
2. Una oración que contiene palabras calificadas como sentimiento puede no contener ninguno. Por ejemplo: “Si encontrara una buena cámara en la tienda, la compraría”.
3. Las oraciones irónicas o sarcásticas con o sin sentimiento son difíciles de detectar. Por ejemplo: “¡Qué rico estaba ese café con sal!”, cuando es muy poco probable que existan amantes de esta combinación.

Por otro lado, y desde el punto de vista del modelado de series temporales, aunque estos modelos ayuden a realizar pronósticos sobre la serie, no son capaces de pronosticar las intervenciones ni tampoco los cambios estructurales producidos por estas.

Los cambios estructurales tienen un efecto muy grande en el modelado de series, haciendo que su correcto análisis sea la parte principal de la caracterización de la dinámica de la serie. Todo ello implica una elevada dificultad en la identificación de los modelos, la estimación y el diagnóstico. Dado que la existencia de cambios genera una alta volatilidad en el error, incluso pueden derivar en patrones en los residuales que lleven a una interpretación errónea del modelo.

Otra limitación de estos métodos o técnicas es su naturaleza subjetiva, a la hora de poder categorizar una relación causa-efecto entre la fuerza exógena y el cambio estructural, la cual los hace vulnerables a la crítica. Para dotar de mayor rigor esta relación J. Bai et al. [52] desarrollan pruebas para la detección del cambio estructural y también del momento exacto de cambio.

A pesar de las limitaciones el análisis de Intervención en particular y las Series temporales en general se muestran como métodos de gran utilidad para extraer gran cantidad de información de la evolución del sentimiento y la opinión pública en la web.

REFERENCIAS

- [1] B. Liu, «Sentiment Analysis and Subjectivity.», *Handbook of natural language processing*, vol. 2, pp. 627–666, 2010.
- [2] H. S. Alim, «What if We Occupied Language?», *The New York Times*, 21-dic-2011.
- [3] B. N. Field y A. Botti, *Politics and Society in Contemporary Spain: From Zapatero to Rajoy*. Palgrave Macmillan, 2013.
- [4] D. Tuminan y K. Becker, «Sentiment-Based Features for Predicting Election Polls: A Case Study on the Brazilian Scenario», en 2014 IEEE/WIC/ACM International Joint Conferences on Web Intelligence

- (WI) and Intelligent Agent Technologies (IAT), 2014, vol. 2, pp. 126-133.
- [5] S. Benabderrahmane, N. Mellouli, M. Lamolle, y P. Paroubek, «Smart4Job: A Big Data Framework for Intelligent Job Offers Broadcasting Using Time Series Forecasting and Semantic Classification», *Big Data Research*, vol. 7, n.o Supplement C, pp. 16-30, mar. 2017.
- [6] L. Bernal, J. A. A. Gasparrini, C. M. Artundo, y M. McKee, «The effect of the late 2000s financial crisis on suicides in Spain: an interrupted time-series analysis», *Eur J Public Health*, vol. 23, n.o 5, pp. 732-736, oct. 2013.
- [7] C. Chen y L.-M. Liu, «Joint Estimation of Model Parameters and Outlier Effects in Time Series», *Journal of the American Statistical Association*, vol. 88, n.o 421, pp. 284-297, mar. 1993.
- [8] J. Fernández-Cruz, «Can Sentiment in our Words be Quantified? An Introduction to Lingmotif, a Sentiment Analysis Software Tool and its Educational Application», *International Congress on the Didactics of the English Language*, vol. 2, n.º 1, jun. 2017.
- [9] J. Fernández-Cruz, «The Expression of Polarity in Financial Discourse and its Application in Sentiment Analysis», Universidad de Málaga, Málaga, 2014.
- [10] Moreno-Ortiz, Antonio, F. Pineda Castillo, y R. Hidalgo García, «Análisis de Valoraciones de Usuario de Hoteles con Sentitext: un sistema de análisis de sentimiento independiente del dominio», *Procesamiento del lenguaje natural*, vol. 45, pp. 31-39, 2010.
- [11] A. Moreno-Ortiz, C. Pérez-Hernández, y R. Hidalgo-García, «Domain-neutral, Linguistically-motivated Sentiment Analysis: a performance evaluation», en *Actas del XXVII Congreso de la SEPLN*, Huelva, Spain, 2011, pp. 847-856.
- [12] A. Moreno-Ortiz y C. P. Hernández, «Lexicon-Based Sentiment Analysis of Twitter Messages in Spanish», *Procesamiento del Lenguaje Natural*, vol. 50, n.º 0, pp. 93-100, dic. 2012.
- [13] A. Moreno-Ortiz, y C. Pérez Hernández. «Lexicon-Based Sentiment Analysis of Twitter Messages in Spanish». En *Procesamiento del Lenguaje Natural, Revista no 50*, marzo de 2013, pp 93-100
- [14] A. Moreno-Ortiz y C. Pérez-Hernández, «Managing Multiword Expressions in a Lexicon-Based Sentiment Analysis System for Spanish», en *Proceedings of the 9th Workshop on Multiword Expressions*, Atlanta, USA, 2013, pp. 1-10.
- [15] A. Moreno-Ortiz y J. Fernández-Cruz, «Identifying polarity in financial texts for sentiment analysis: a corpus-based approach», *Procedia-Social and Behavioral Sciences*, vol. 198, pp. 330-338, 2015.
- [16] A. Moreno-Ortiz, y C. Gómez Pascual, «Extracting Domain-Specific Features for Sentiment Analysis Using Simple NLP Techniques: Running Shoes Reviews», *EPIC Series in Language and Linguistics*. 2016, pp. 298-307
- [17] A. Moreno-Ortiz y C. Pérez-Hernández, «Tecnología Lingmotif at TASS 2017: Spanish Twitter Dataset Classification Combining Wide-coverage Lexical Resources and Text Features», en TASS 2017: Workshop on Semantic Analysis at SEPLN, Murcia, Spain, 2017, pp. 35-42.
- [18] A. Moreno-Ortiz, «Lingmotif: A User-focused Sentiment Analysis Tool», *Procesamiento de Lenguaje Natural*, vol. 58, 2017.
- [19] A. Moreno-Ortiz, «Lingmotif: Sentiment Analysis for the Digital Humanities», presentado en Proceedings of the EAACL, Valencia, Spain, 2017.
- [20] J. Lyons, *Semantics (vols I & II)*. Cambridge, UK: CUP, 1977.
- [21] N. Besnier, «Reported speech and affect on Nukulaeae Atoll», *Responsibility and evidence in oral discourse*, pp. 161-181, 1993.
- [22] M. A. K. Halliday y C. M. I. M. Matthiessen, *Halliday's Introduction to Functional Grammar*. Routledge, 2013.
- [23] S. Hunston y G. Thompson, Eds., *Evaluation in text: authorial stance and the construction of discourse*, Reprinted. Oxford: Oxford Univ. Press, 2003.
- [24] J. Martin y P. White, *The Language of Evaluation: Appraisal in English*, Edición: 2005. Basingstoke, Hampshire: Palgrave Macmillan, 2008.
- [25] M. Bednarek, *Evaluation in media discourse: analysis of a newspaper corpus*. London: Continuum, 2008.
- [26] M. Bednarek, *Emotion talk across corpora*. Basingstoke [England]; New York: Palgrave Macmillan, 2008.
- [27] J. G. Carbonell, «Subjective Understanding: Computer Models of Belief Systems.», Yale University, New Haven, 1979.
- [28] B. Liu, «Sentiment Analysis and Opinion Mining», *Synthesis Lectures on Human Language Technologies*, vol. 5, n.º 1, pp. 1-167, may 2012.
- [29] B. Liu, *Sentiment analysis: mining opinions, sentiments, and emotions*. New York, NY: Cambridge University Press, 2015
- [30] A. B. Pawar, M. A. Jawale, y D. N. Kyatanavar, «Fundamentals of Sentiment Analysis: Concepts and Methodology», en *Sentiment Analysis and Ontology Engineering*, vol. 639, W. Pedrycz y S.-M. Chen, Eds. Cham: Springer International Publishing, 2016, pp. 25-48.
- [31] A. Moreno-Ortiz, *Lingmotif*. Málaga (Spain): Universidad de Málaga, 2016.
- [32] Y. Jo y A. H. Oh, «Aspect and sentiment unification model for online review analysis», en Proceedings of the fourth ACM international conference on Web search and data mining, 2011, pp. 815-824.
- [33] B. Pang y L. Lee, «Opinion mining and sentiment analysis», *Foundations and trends in information retrieval*, vol. 2, n.º 1-2, pp. 1-135, 2008.
- [34] P. D. Turney, «Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews», en *Proceedings of the 40th annual meeting on association for computational linguistics*, 2002, pp. 417-424.
- [35] B. Pang, L. Lee, y S. Vaithyanathan, «Thumbs up?: sentiment classification using machine learning techniques», en *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, 2002, pp. 79-86.
- [36] B. Pang y L. Lee, «Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales», *arXiv:cs/0506075*, jun. 2005.
- [37] L. Polanyi y A. Zaenen, «Contextual Valence Shifters», en *Computing Attitude and Affect in Text: Theory and Applications*, J. G. Shanahan, Y. Qu, y J. Wiebe, Eds. Springer Netherlands, 2006, pp. 1-10.
- [38] A. Andreevskaia y S. Bergler, «When Specialists and Generalists Work Together: Overcoming Domain Dependence in Sentiment Tagging.», 2008.
- [39] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, y M. Stede, «Lexicon-based methods for sentiment analysis», *Computational linguistics*, vol. 37, n.º 2, pp. 267-307, 2011.
- [40] A. Esuli y F. Sebastiani, «SENTIWORDNET: A high-coverage lexical resource for opinion mining», Technical Report 2007-TR-02, Istituto di Scienza e Tecnologie dell'Informazione, Consiglio Nazionale delle Ricerche, Pisa, IT, 2007.
- [41] A. Montejo-Raez, E. Martínez-Cámara, M. T. Martín-Valdivia, y L. A. Urena-Lopez, «Random walk weighting over sentiwordnet for sentiment polarity detection on twitter», en *Proceedings of the 3rd Workshop in Computational Approaches to Subjectivity and Sentiment Analysis*, 2012, pp. 3-10.
- [42] B. Ohana y B. Tierney, «Sentiment classification of reviews using SentiWordNet», en *9th. IT & T Conference*, 2009, p. 13.
- [43] B. Varga y A. Groza, «Integrating dbpedia and sentiwordnet for a tourism recommender system», en *Intelligent Computer Communication and Processing (ICCP), 2011 IEEE International Conference on*, 2011, pp. 133-136.
- [44] M. Marshall y J. Catlin, «Lexalytics», *Lexalytics*, 2016. [En línea]. Disponible en: <https://www.lexalytics.com/about>.
- [45] N. S. Balke, «Detecting Level Shifts in Time Series», *Journal of Business & Economic Statistics*, vol. 11, n.o 1, pp. 81-92, ene. 1993.
- [46] P. Perron, «The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis», *Econometrica*, vol. 57, n.º 6, pp. 1361-1401, 1989.
- [47] J. Bai, «Estimation of a Change Point in Multiple Regression Models | The Review of Economics and Statistics | MIT Press Journals».
- [48] A. C. Atkinson, S. J. Koopman, y N. Shephard, «Detecting shocks: Outliers and breaks in time series», *Journal of Econometrics*, vol. 80, n.º 2, pp. 387-422, oct. 1997.
- [49] D. I. Harvey, S. J. Leybourne, y P. Newbold, «Seasonal unit root tests with seasonal mean shifts», *Economics Letters*, vol. 76, n.º 2, pp. 295-302, jul. 2002.

- [50] R. Kaiser, y A. Maravall, «Seasonal outliers in time series», Journal of the InterAmerican Statistical Institute, special issue on time series, 2002.
- [51] P. Blinov, M. Klekovkina, E. Kotelnikov, y O. Pestov, «Research of lexical approach and machine learning methods for sentiment analysis», *Computational Linguistics and Intellectual Technologies*, vol. 2, n.º 12, pp. 48–58, 2013.
- [52] J. Bai, «Critical values for multiple structural change tests» The Econometrics Journal - Wiley Online Library, 2003.

Incidencia de la comunicación visual en la evaluación heurística

Marmolejo-Cueva, Melba-Cristina¹, Grob, Marc²

melba.marmolejo@pucese.edu.ec, marc.grob@pucese.edu.ec

¹Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador, Escuela de Diseño Gráfico

²Pontificia Universidad Católica del Ecuador-Sede Esmeraldas- Ecuador, Escuela de Sistema y Computación

Resumen— La Evaluación Heurística es una técnica que permite establecer deficiencias de uso durante la interacción con plataformas digitales como es el caso de las páginas web. En este artículo se presenta el contraste de criterios entre dos especialistas en las áreas de sistemas tecnológicos y el diseño gráfico respecto a la incidencia de la comunicación visual en la evaluación heurística a partir de un estudio descriptivo, documental cuantitativo-cualitativo de la usabilidad en cuatro plataformas universitarias seleccionadas bajo el criterio de ofertar la carrera de Diseño Gráfico en la Zona 1 de Ecuador (que comprende a Imbabura, Carchi, Esmeraldas y Sucumbíos) durante el año 2017. Para la recolección de datos se empleó una lista de cotejo basada en la teoría [1] con enfoque visual, implementada por cinco expertos en áreas como el diseño gráfico, marketing y educación de la Pontificia Universidad Católica del Ecuador Sede Esmeraldas PUCESE. El análisis de datos se estructuró en tres fases de comparación: percepción de uso, detección de falencias y la Evaluación Heurística con enfoque visual. La investigación permitió evidenciar que la orientación visual de la Evaluación Heurística permite reconocer aspectos puntuales sobre la adecuada organización de elementos que integran la plataforma, como cromática, tipografía y distribución en la composición, facilitando la usabilidad al determinar el grado de satisfacción del usuario a nivel de interfaz. Finalmente se hace una reflexión desde dos perspectivas (sistemas tecnológicos y diseño gráfico) acerca del aporte de la comunicación visual en el desarrollo de soluciones digitales, trascendiendo de la estética a la funcionalidad e incidiendo en el nivel de usabilidad del sistema tecnológico.

Palabras clave: Evaluación Heurística, Comunicación Visual, Usabilidad

I. INTRODUCCIÓN

Los nuevos modelos de comunicación apoyados por las Tecnologías de la Información y Comunicación TIC, se caracterizan por permitir, a grandes rasgos, el acceso a información y en este sentido, las plataformas digitales se han convertido en algo más que un medio de intercambio de datos, al revolucionar la forma de interactuar con la información generada por usuarios con un rol activo. Los equipos de trabajo encargados del desarrollo de proyectos tecnológicos generalmente están relacionados con las competencias informáticas, conocen los procesos, mecanismos, herramientas y normas para la generación de plataformas digitales como los sitios web. Entre las etapas de desarrollo se encuentra la validación de prototipos con el objetivo de establecer si se cumple eficientemente con las tareas planteadas.

[2] define a la Evaluación Heurística como un método para valorar deficiencias de uso dentro del proceso iterativo de diseño de interfaces de usuario, formando parte de la ingeniería

de usabilidad, disciplina definida por [3] como el trabajo de planificación, implementación y verificación de los objetivos de uso de un sistema. Además, expresan que la usabilidad Web se originó con la comunicación en la internet, llegando a la cúspide de su desarrollo gracias al aporte de [4], cuyas investigaciones sentaron las bases teórico-prácticas para lograr que las aplicaciones sean funcionales y por ende faciliten la interacción del usuario. La importancia de la usabilidad radica en la comodidad percibida al usar una herramienta específica. Dichas características se consiguen a través de la atracción y calidad durante el uso, lo que constituye el nuevo paradigma originado desde la funcionalidad hasta la facilitación de trabajo del usuario.

Pueden cumplir el rol de evaluadores, profesionales de las ciencias informáticas como programadores, estudiantes o con habilidades afines al ramo. En relación con otros métodos de evaluación de usabilidad, la Heurística es más conveniente si se cuenta con expertos del área y suele ser más económico con respecto a otros métodos, lo que justifica su alta demanda. Este método permite que las instrucciones para el uso del sistema sean visibles y accesibles para que el evaluador tenga la capacidad de detectar y resolver posibles deficiencias que afecten la interacción de los usuarios ya sean estos inexpertos o experimentados.

Con la evolución de los modelos comunicacionales, las teorías de la comunicación orientadas hacia lo visual cobran protagonismo. En este orden de las cosas, [5] hace una reflexión profunda sobre el diseño de comunicación visual y de la persona como su razón de ser (receptor). Lo presenta como un epifenómeno que avanza a la par con la tecnología, textualmente expresa que: *“Diseño de comunicación visual es una disciplina transversal de información. Él, aporta asimismo su lenguaje específico a las demás disciplinas de diseño. Su objeto es hacer el entorno más inteligible y mejor utilizable, contribuir a la calidad de vida de las personas y a la sociedad del conocimiento”*. Entre los tópicos que el autor menciona como misión del diseño de comunicación visual están:

- Reducir la complejidad del entorno y de los procesos.
- Hacer el entorno más inteligible, agradable y mejor utilizable.
- Incorporar información útil (semántica y estética) al entorno, a las construcciones y a los objetos para optimizar su uso o su disfrute.
- Adaptar la estrategia comunicacional y el lenguaje visual a la cultura media del grupo destinatario, a sus códigos generacionales o profesionales y a sus expectativas.

- Organizar la coherencia estética, formal y significativa de los mensajes y las informaciones.
- Diseñar interfaces, servicios y modelos que faciliten sus prestaciones a los usuarios.
- Visualizar las realidades invisibles para hacerlas comprensibles y memorizables, incorporándose así a la cultura personal.
- Generalmente los estudios sobre usabilidad están encaminados a evaluar la rapidez de respuesta de las plataformas ante la acción del usuario sin considerar que la parte visual como un aspecto relevante en acciones como la memorización y desempeño eficaz.

El tipo de información que se propaga en las redes tecnológicas posee alto grado de complejidad debido al papel dinámico de los usuarios que interactúan bidireccionalmente, pasando de un rol pasivo como receptores a otro más activo como difusores. Este cambio afecta la forma de acceso a la información, modificando la simplicidad de sus funciones.

Los equipos de desarrollo deben estar conformados multidisciplinariamente, con profesionales de campos transversales como los diseñadores gráficos y comunicadores visuales. De esta preocupación, surge la necesidad de contrastar criterios desde dos ópticas para establecer si la particularización de la teoría visual en la Evaluación Heurística aporta a la usabilidad de las páginas web.

II. DESARROLLO

1. Metodología

Para este artículo, dos especialistas de las áreas de ingeniería en sistemas y diseño gráfico han realizado una lectura crítica de la investigación de [6] que presenta una adaptación de la Evaluación Heurística enfatizando en los componentes visuales. Para ello, se seleccionó como objeto de estudio a las páginas web de las universidades localizadas en la zona 1 de Ecuador (Imbabura, Carchi, Esmeraldas y Sucumbíos) que ofertan la carrera de diseño gráfico. A continuación, se creó una lista de cotejo con indicadores relacionados a la comunicación visual y bajo la teoría de [1], organizando aspectos de dicha teoría en cuatro categorías: Contenido de información, Contenido visual, Navegación o Interactividad y Presentación. Cada aspecto contenía sus propios indicadores con una escala de evaluación mixta (estándar y numérica), cuyos rangos de valor se estructuraban así: Totalmente de Acuerdo (TD=5), De acuerdo (D=4), Indiferente (ID=3), Desacuerdo (DS=2) y Totalmente Desacuerdo, (TDS=1). Este instrumento (lista de cotejo) fue aplicada por evaluadores seleccionados a partir de docentes de la carrera de Diseño Gráfico que pertenecen a diversas disciplinas como la educación, diseño y tecnología.

Para llevar los datos cualitativos a datos cuantitativos se aplicó una fórmula “Regla de tres” cuantificando el número de veces que se seleccionó una escala y luego dividiendo para la cantidad de ítems evaluados entre todas las universidades.

acceso a la información por cansancio visual u obstrucción de lectura.

El análisis de los datos consta de tres fases de comparación: la primera concierne a la percepción de uso para determinar la escala y el grado de aceptación de cada web universitaria; la segunda fase compara los ítems por cada universidad, verificando posibles deficiencias en el uso de las plataformas y la tercera fase se realizó a la evaluación heurística adaptada a lo visual.

Posteriormente se realiza una lectura crítica de los resultados desde dos puntos de vista: el informático y del diseño gráfico para desarrollar una discusión que contraste ambas percepciones y conocer las posibilidades de coexistencia de las disciplinas señaladas.

III. RESULTADOS

En la primera fase de estudio, el equipo de evaluadores, según la escala propuesta, estuvo de acuerdo con el nivel de usabilidad de los sitios investigados, aunque con observaciones de mejora y recomendaciones. En la segunda fase se hizo el análisis comparativo entre los ítems de las cuatro categorías y las percepciones de los evaluadores obteniendo los siguientes resultados:

Categoría Contenido de Información. – En cuanto al contenido textual de los sitios web, la presentación visual se percibe como agradable para el usuario y es de fácil entendimiento, no sólo por la correcta selección tipográfica sino por el estilo claro y breve en la redacción del mensaje, lo que previene el cansancio visual, así como la distracción durante el proceso de interacción con el contenido.

Categoría Contenido Visual. - Consta de cinco aspectos que miden la calidad en la usabilidad enmarcados en la estructura y organización de los contenidos visuales a través de la percepción del usuario. Este es uno de los aspectos relevantes para la adaptación del modelo heurístico, puesto que influye en el desenvolvimiento intuitivo y fluidez de acción sobre la interfaz. En el caso de los íconos, deben poseer una línea gráfica porque aporta orden a la composición y facilita el reconocimiento de cada una de las funciones. Estéticamente, los íconos bien diseñados contribuyen a la comunicación rápida y precisa, además, se comprueba que las deficiencias de organización en íconos, imágenes, botones, cajas de texto, titulares, y elementos multimedia, pueden afectar tareas fundamentales como la lectura o búsqueda de información. También se otorga control al usuario disminuyendo la cantidad de elementos gráficos lo que incrementa su dominio del sistema. Los mensajes de pantalla visualmente bien diseñados facilitan la percepción del mensaje, y dotan al usuario de un estado de seguridad al presentarse un problema. Cuando se evalúan aspectos de la interfaz como los colores, contenidos de texto e imagen o ventanas de navegación, desde el enfoque visual, se proporciona información especializada sobre comunicación para los desarrolladores que eviten problemas de

Categoría Navegación o Interactividad. - La navegabilidad depende de la interacción entre el software y el usuario. La adaptación de la evaluación permitió confirmar que el diseño visual y la composición de los elementos de búsqueda, favorecen a la navegación.

Categoría Presentación. – Esta categoría complementa las valoraciones recogidas en la categoría de contenido visual a través de la cromática, contenido textual, así como visual entre otros componentes que permitan facilidad de uso en la plataforma.

IV. DISCUSIÓN

Aunque la Evaluación Heurística es una herramienta concebida desde la óptica informática, la adaptación de la teoría de comunicación visual enriquece aspectos considerados como esenciales para la interacción persona-dispositivo porque promueve el uso intuitivo por medio de atributos de la comunicación que sólo pueden materializarse con el diseño, entre los que se puede mencionar al minimalismo (que propende a la claridad del mensaje, sea este textual o gráfico), la composición y la jerarquización de los elementos que lo integran.

Se afirma [7] que una estrategia eficaz de comunicación en la web que permite validar su calidad es el diseño visual y la arquitectura de la información, siendo indicadores en el primer caso la iconicidad, el cromatismo y la lingüística. Significa entonces que ya se ha contemplado factores visuales como referencia de calidad, aunque, por otro lado, [8] expresa que, de las diferentes opciones de evaluación se destaca el estándar ISO/IEC 9126 de evaluación externa, (basado principalmente en *Funcionalidad y Calidad de uso*), por ser una de las más efectivas aplicadas a las herramientas web y en dicho modelo no se priorizan indicadores relacionados con las teorías de la comunicación visual.

El especialista del área informática sostiene que el enfoque visual contribuye relativamente, debido a que sólo enfatiza algunos aspectos que ya están considerados en la heurística y aunque reconoce la importancia de la interfaz dentro del proceso de interacción, ésta constituye sólo una parte de todo lo que constituye una plataforma.

Por otro lado, el especialista del área de diseño gráfico aluce que, al resaltar los indicadores heurísticos más orientados a teorías de la comunicación visual, se hacen más perceptibles las disfuncionalidades porque la interfaz es la primera cara con la que se enfrenta el usuario para hacer su trabajo. Desde la informática, la detección integral de las deficiencias no puede enfocarse principalmente en la capa de la interfaz, porque, aunque es el espacio de interacción directa usuario-sistema, están otros aspectos relativos a otras capas más profundas que integran al sistema, no obstante, la perspectiva de comunicación visual reconoce que las plataformas digitales, entre éstas la web, están constituidas por un sistema complejo pero que es precisamente la interfaz el primer acercamiento entre el usuario y las funciones, por lo tanto,

aunque no es el único componente, es el primero que da la posibilidad para detectar problemas iniciales que posteriormente revelarán deficiencias más profundas.

La informática reconoce que su campo de acción busca resolver tareas a través de sistemas que permiten la automatización de procesos, pero su competencia no profundiza en las estrategias para comunicar eficientemente las funciones para las que fue creada. Conociendo esta falencia, la participación del diseño gráfico deja de ser complementaria y trasciende de la estética para convertirse en otro punto clave en la funcionalidad del sistema. Bajo este criterio, se puede considerar el caso [9] que presenta un modelo heurístico que consta de arquitectura, diseño y visualización de información médica, su principal aporte está relacionado con la temporalidad y ubicuidad en los procesos culturales, comerciales y/o educativos.

[10] Para el análisis del mensaje visual, más allá del medio de aplicación, debe considerarse la estrategia de representación visual generada como producto de la composición o combinación del color, cesía, forma, textura y movimiento. Estos conceptos, son claros para los especialistas en el área del diseño gráfico, que por su naturaleza transversal aporta significativamente a proyectos diversa índole. En esta misma línea de pensamiento, el propósito del comunicador visual dentro del equipo de desarrollo es representar gráficamente conceptos abstractos originados en la programación para que el usuario pueda identificar la funciones que la plataforma le ofrece y luego reconocer lo que el sistema necesita antes, durante y después de cada acción

V. CONCLUSIONES

1. La Evaluación Heurística es un método de validación de la usabilidad de los sistemas digitales que ha probado su efectividad y que puede incrementar su potencialidad para detectar deficiencias de funcionalidad a través de la implementación de las teorías de la comunicación visual.
2. El estudio de [6] evidencia la viabilidad de una Evaluación Heurística adaptada a los conceptos de comunicación visual, demostrando que los indicadores relacionados a los componentes visuales, si aportan en la detección de problemas en la funcionalidad de los sitios web evaluados
3. Los especialistas de las áreas informáticas y de diseño gráfico presentaron inicialmente puntos de vista divergentes en el grado de incidencia de la comunicación virtual para valorar la usabilidad en plataformas digitales. Por un lado, se reconoce un aporte relativo del diseño gráfico en la capa de interfaz, que constituye sólo una parte de todo el conjunto del sistema, mientras que, desde otra perspectiva la comunicación visual trasciende de la estética y ayuda en la detección de problemas de uso por ser la primera “cara” de la plataforma digital.
4. Finalmente, los especialistas concuerdan en que la comunicación visual permite representar conceptos

abstractos originados en las fases más profundas dentro del proceso de desarrollo de un sistema facilitando la interacción y por ende la usabilidad.

5. Es necesario que los equipos de desarrollo sean multidisciplinarios, ya que el aporte de un comunicador visual va más allá de la estética o la facilidad de uso de una plataforma tecnológica, permite visualizar funciones y este hecho hace más intuitivo a un sistema.

REFERENCIAS

- [1] J. Nielsen y R. Molich, «HEURISTIC EVALUATION OF USER INTERFACES,» *CHI'90 Proceedings*, pp. 249-256, Abril 1990.
- [2] W. Sánchez, «La usabilidad en Ingeniería de Software: definición y características,» *Ing-Novación. Revista de Ingeniería e Innovación de la Facultad de Ingeniería. Universidad Don Bosco*, n° 1, pp. 7-21, Agosto 2011.
- [3] L. Perurena Cancio y M. Moráguez Bergues, «Usabilidad de los sitios Web, los métodos y las técnicas para la evaluación,» *Revista Cubana de Información en Ciencias de la Salud*, vol. 24, n° 2, abril-junio 2013.
- [4] J. Nielsen, *Technical University of Denmark. Department of Computer Science. DK-2800 Lyngby Copenhagen Denmark. dat JN@NEUVM1 . bitnet*, 1990.
- [5] J. Costa, «Diseño de Comunicación Visual: El nuevo paradigma,» *Gráfica*, vol. 2, n° 4, pp. 89-107, 2014.
- [6] C. Quinto, «EVALUACIÓN HEURÍSTICA CON ENFOQUE DE COMUNICACIÓN VISUAL EN SITIOS WEB DE LAS UNIVERSIDADES QUE OFERTEN LA CARRERA DE DISEÑO GRÁFICO EN LA ZONA 1 DEL ECUADOR AÑO,» Esmeralda, 2017.
- [7] D. Grávalos Macho, «La calidad de una página web como herramienta de comunicación,» *Estudios sobre el mensaje peridístico*, vol. 19, n° especial marzo, pp. 253-261, 2013.
- [8] C. X. Vega Oyola, J. K. Celleri Pacheco, J. L. Maza-Córdova y K. O. Sarmiento Chugcho, «Validación de herramientas on line que miden la calidad de sitios web,» de *Conference Proceedings 2017. Universidad Técnica de Machala*, Machala, 2017.
- [9] G. I. Garmendia Ramírez, «Modelo heurístico para la visualización de información médica,» *Zaloamati*, pp. 1-16, 2014.
- [10] J. L. Caivano, «Semiótica, cognición y comunicación visual: los signos básicos que construyen lo visible,» *Tópicos del Seminario. Revista de semiótica*, vol. 1, n° 13, pp. 113-135, 2015.

Determinación del ritmo cardíaco en tiempo real para el diagnóstico de afecciones relacionadas al corazón mediante la aplicación de redes neuronales.

Bryan Rivas Gracia¹, Brian Robles Bernal¹, Marcelo Sánchez Tarira¹, Franyelit Suárez¹
 bryan.rivas@pucese.edu.ec, brian.robles@pucese.edu.ec, marcelo.sanchez@pucese.edu.ec, franyelit.suarez@pucese.edu.ec
¹Pontificia Universidad Católica del Ecuador - Sede Esmeraldas – Ecuador

Resumen—Este documento contiene los datos acerca de una investigación realizada sobre la aplicación de las redes neuronales en la determinación, mediante un dispositivo electrónico capaz de medir en tiempo real la frecuencia cardíaca y posición del usuario para diagnosticar posibles afecciones relacionadas al corazón.

Palabras claves—redes neuronales; percepción; backpropagation; ritmo cardíaco

I. INTRODUCCIÓN

Cuando se habla de un órgano tan importante como es el corazón, se trata de aquel órgano o aquel musculo que se encarga de bombear la sangre a través de los vasos sanguíneos del cuerpo, por medio de los vasos sanguíneos eritrocitos permite que se distribuya por todo el cuerpo, alimentando con oxígeno las células que forman los tejidos. [2] Un fallo en este órgano puede tener consecuencias fatales. Sin oxígeno, los tejidos del cuerpo humano sufrirán una isquemia, en esta situación puede que las células dejen de cumplir sus funciones debido a la incapacidad de mantener su metabolismo. Si la situación se prolonga, las células podrían morir dando lugar a una necrosis. Las consecuencias de las isquemias y necrosis son amplias y los efectos de estas varían según la zona en la que se den.

La pulsera podrá leer el pulso cardíaco de una determinada persona y en base a esto se podrá determinar el estado del ritmo cardíaco de la misma, dependiendo del pulso q se tenga se detectará las posibles dolencias o enfermedades. [3] manifiesta La frecuencia cardíaca o ritmo cardíaco es el número de latidos por unidad de tiempo, estos generalmente se miden en cantidad de pulsaciones por minuto, en seres humanos, la medida del ritmo cardíaco normal se encuentra entre 70 y 90 latidos o pulsaciones por minuto. Además, esta medición se utiliza para ayudar a diagnosticar enfermedades cardíacas y para el seguimiento de distintas condiciones médicas.

La pulsera indicara la lectura de la persona y de ser necesario el lugar de atención medica más cercano, por esto deberá incorporar una bocina y GPS además del lector de pulso; además se contará con una aplicación web en la que constaran los datos de la persona para un control y localización en tiempo real para indicar los lugares de atención medica más cercano.

Para las mediciones de los respectivos datos se tendrán en cuenta los signos vitales q presenten cada persona. [4] Los signos vitales (SV) son la temperatura (T), Frecuencia Cardíaca pulso (P), la respiración (R) y la presión sanguínea

(PS) de una persona. El dolor, es también denominado el quinto signo vital. El estado de salud de un individuo se refleja en estos indicadores de la función corporal. Un cambio en los signos vitales puede identificar una modificación en el estado de la salud.

VARIACIONES NORMALES DE LOS SIGNOS VITALES RELACIONADAS CON LA EDAD				
Edad	Temperatura	Pulso (lat./min)	Respiración (resp./min)	Presión Sanguínea (mm Hg)
Recién nacido	36.8 C (98.2 F) (axilar)	80-180	30-60	73/55
1-3 años	37.7 C (99.9 F) (rectal)	80-140	20-40	90/55
6-8 años	37 C (98.6 F) (bucal)	75-120	15-25	95/75
10 años	37 C (98.6 F) (bucal)	75-110	15-25	102/62
Adolescentes	37 C (98.6 F) (bucal)	60-100	15-20	102/80
Adultos	37 C (98.6 F) (bucal)	60-100	12-20	120/80
>70 años	37 C (98.6 F) (bucal)	60-100	15-20	120/80

Fig. 1. Valores de signos vitales según la edad [4]

Muchas son las personas que piensan q se encuentran de excelente forma, si saber cómo se encuentra en realidad su estado de salud y mediante la realización de esta pulse ayudara a conocer el estado en la que se encuentra. [5] A pesar de que la hipertensión se puede controlar y tratar de forma efectiva mediante fármacos, una cuarta parte de las personas que la padecen ni siquiera lo sabe. Además, de entre quienes sí conocen que su presión arterial es elevada, sólo dos tercios reciben tratamiento y menos de la mitad la tienen bajo control.

Se brinda una gran ayuda a la sociedad, ya que no es necesario tener conocimientos médicos, sino ofrecer una ayuda por medio de la tecnología que actualmente se tiene. [6] Se ha hecho hincapié en las metodologías de trabajo así como en las tecnologías usadas, problemas encontrados y tomas de decisiones a las que se ha enfrentado durante el desarrollo en estas nuevas tecnologías usadas para el desarrollo de una aplicación orientada al campo de la salud que se encarga de realizar mediciones de forma pasiva sin dar conocimiento al usuario sobre lo mismo y pensada para después de la monitorización, ofrecer un informe sobre el estado del usuario paciente.

II. REDES NEURONALES

Las redes neuronales es una de la base más importante para el desarrollo de la Inteligencia Artificial (IA), tienen como inspiración copiar o simular el comportamiento del cerebro humano, permitiendo crear sistemas capaces de

aprender y solucionar problemas de gran dificultad. Se puede decir que estas redes son más que una simple emulación de algunas características originarias de los humanos, ya que posee la capacidad de aprender y relacionar hechos.

Existen problemas que no basta con un algoritmo para poder darle una solución, sino que para llegar a esta se hace uso de la experiencia, ya que el hombre es capaz de solucionar problemas de acuerdo a la experiencia que este tenga. [7] En definitiva, las redes neuronales no son más que un modelo artificial y simplificado del cerebro humano, que es el ejemplo más perfecto del que disponemos para un sistema que es capaz de adquirir conocimiento a través de la experiencia. Una red neuronal es “Un nuevo sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona”.

a. Perceptrón

El perceptrón es la neurona más simple en este ámbito, consiste en la función de una activación binaria que arroja dos posibles resultados. Está constituido de una sola capa siendo una red feed-forward, la cual se basa en la transferencia de umbral.

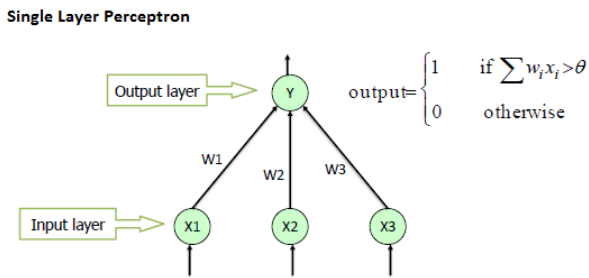


Fig. 2. Diagrama de Perceptrón Simple

[8]El perceptrón de una sola capa no tiene un conocimiento a priori, por lo que los pesos iniciales se asignan aleatoriamente. SLP suma todas las entradas ponderadas y si la suma está por encima del umbral (un valor predeterminado), se dice que SLP está activada (salida = 1).

OUTPUT

$$1. W_1X_1 + W_2X_2 + \dots + W_nX_n > 0 \rightarrow 1$$

$$2. W_1X_1 + W_2X_2 + \dots + W_nX_n \leq 0 \rightarrow 0$$

Pesos del Perceptrón Simple

Los valores de entrada se presentan al perceptrón, y si la salida pronosticada es la misma que la salida deseada, entonces el rendimiento se considera satisfactorio y no se realizan cambios en los pesos. Sin embargo, si la salida no coincide con la salida deseada, entonces los pesos deben cambiarse para reducir el error.

$$3. \Delta w = n \times d \times x$$

$d \rightarrow$ Predicted output – desired output

$n \rightarrow$ Learning rate, usually less than 1

$x \rightarrow$ Input data

Ajuste de Pesos del Perceptrón Simple

Como SLP es un clasificador lineal y si los casos no son separables linealmente, el proceso de aprendizaje nunca llegará a un punto en el que todos los casos se clasifiquen correctamente. El ejemplo más famoso de la incapacidad del perceptrón para resolver problemas con casos linealmente no separables es el problema XOR.

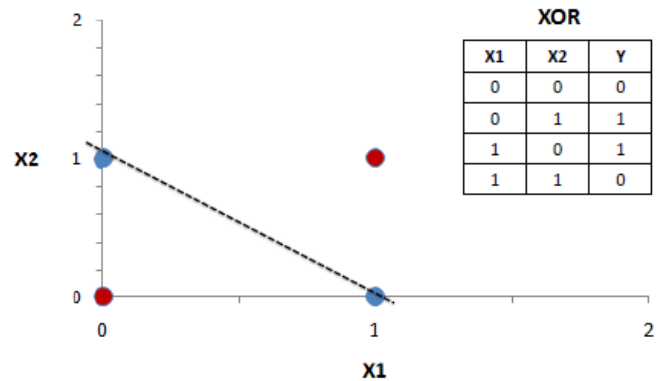


Fig. 3. Ejemplo del problema XOR

b. Aprendizaje de Tipo BackPropagation

Básicamente es un algoritmo de aprendizaje donde dada una red neuronal artificial y una función de error, el método calcula el gradiente de la función de error con respecto a los pesos de la red neuronal. Es una generalización de la regla delta para perceptrones a redes neuronales de alimentación directa multicapa.

La popularidad de BackPropagation ha experimentado un resurgimiento reciente dado la adopción generalizada de redes neuronales profundas para el reconocimiento de imágenes y reconocimiento de voz. Se considera un algoritmo eficiente y las implementaciones modernas aprovechan las GPU especializadas para mejorar aún más el rendimiento.

[9] Tiene como objetivo decidir cómo aprender buenas representaciones internas, es decir, cuáles deberían ser los pesos y los sesgos para los nodos de capa oculta. A diferencia del perceptrón, que tiene la regla delta para aproximar una salida de objetivo bien definida, los nodos de capa oculta no tienen un resultado objetivo, ya que se usan como pasos intermedios en el cálculo.

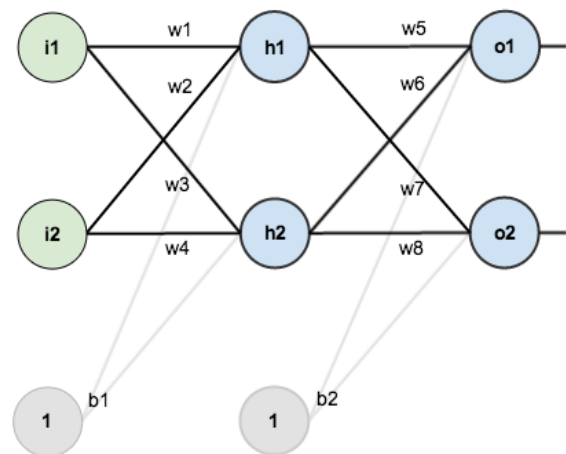


Fig. 4. Estructura Básica del Aprendizaje BackPropagation

[10] Las redes Backpropagation tienen un método de entrenamiento supervisado. A la red se le presenta parejas de patrones, un patrón de entrada emparejado con un patrón de salida deseada. Por cada presentación los pesos son ajustados de forma que disminuya el error entre la salida deseada y la respuesta de la red.

El algoritmo consiste en:

- Calcule la fase de avance para cada par de entrada-salida y almacene los resultados, y para cada nodo en la capa, procediendo desde la capa de entrada, hasta la capa de salida.

$$4. (\vec{X}_d, Y_d) \vec{Y}_d a_j^k o_j^k j k 0 m$$

- Calcule la fase hacia atrás para cada par de entrada-salida y almacene los resultados para cada nodo de conexión de peso en capa en nodo en capa al pasar de la capa de salida, a la capa de entrada.

$$5. (\vec{X}_d, Y_d) \frac{\partial E_d}{\partial w_{ij}^k} w_{ij}^k i k - 1 j k m 1$$

- Evalúa el término de error para la capa final usando la segunda ecuación. b) Vuelva a propagar los términos de error para las capas ocultas, trabajando hacia atrás desde la última capa oculta, utilizando repetidamente la tercera ecuación. c) Evaluar las derivadas parciales del error individual con respecto a usando la primera ecuación δ_1^m

$$6. \delta_j^k k = m - 1$$

$$7. E_d w_{ij}^k$$

- Combine los gradientes individuales para cada par de entrada-salida para obtener el gradiente total para todo el conjunto de pares de entrada-salida usando la cuarta ecuación (un promedio simple de los gradientes individuales).

$$8. \frac{\partial E_d}{\partial w_{ij}^k} \frac{\partial E(X, 0)}{\partial w_{ij}^k} X = \{(\vec{X}_1, Y_1), \dots, (\vec{X}_n, Y_n)\}$$

- Actualice los pesos de acuerdo con la tasa de aprendizaje y el gradiente total usando la quinta ecuación (moviéndose en la dirección del gradiente negativo) α .

$$9. \frac{\partial E(X, 0)}{\partial w_{ij}^k}$$

III. ANTECEDENTES MEDICOS

Para el correcto diagnóstico de una enfermedad cardíaca hacia un paciente, es necesaria la aplicación de diferentes pruebas que garanticen el estado en el que se encuentra el corazón y sus componentes. Según [11], ninguna prueba es capaz de detectar una enfermedad cardíaca por sí sola, por ello, es necesaria la realización de distintas pruebas tales

como: Electrocardiogramas, pruebas de esfuerzo, ecocardiograma, radiografía de tórax, entre otras. Adicional a lo ya anteriormente mencionado, la consideración de la historia clínica y el antecedente familiar del paciente suelen ser puntos claves a la hora de determinar si un paciente es propenso o no a padecer alguna enfermedad de este tipo para su posterior detección y tratamiento.

Para la realización de esta investigación y posterior aplicación al proyecto a desarrollarse, se han considerado 3 aspectos necesarios.

a. Ritmo Cardíaco

[12] afirma que el ritmo cardíaco, es el número de latidos del corazón por la unidad del tiempo y se expresa generalmente como BPM (Beats per Minute) o PPM (Pulsos Por Minuto).

Los latidos ayudan al corazón a limpiar la sangre y oxigenar el cuerpo, además de proveer de sangre limpia al ventrículo izquierdo y a los vasos sanguíneos de la arteria aorta.

La medición de los latidos cardíacos ayuda al diagnóstico de posibles enfermedades del corazón, [13] afirma que un PPM de una persona normal varía entre 70-90. Para la medición de los PPM es necesario tomarle el pulso a la persona, de esta forma se podrá saber cuántas pulsaciones da su corazón en un minuto.

Algunos de los sitios del pulso incluyen:

- En la sien (temporal)
- En el cuello (carotídeo)
- Parte interna del brazo (humeral)
- En la muñeca (radial)
- Parte interna del pliegue del codo (cubital)
- En la ingle (femoral)
- En el dorso del pie (pedio)
- En la tetilla izquierda de bebés (pulso apical)
- Los más comunes son el pulso radial y el carotídeo.

b. Enfermedades que influyen en la Frecuencia Cardíaca

Según [14] los trastornos o enfermedades de influencia cardíaca, consisten en latidos del corazón que van a destiempo, alterando las señales eléctricas del corazón, haciendo que lata muy rápido, o muy despacio.

Según [14], los tipos más importantes de trastornos del ritmo cardíaco incluyen:

- Taquicardia: Un ritmo cardíaco rápido (más de 100 latidos por minuto)
- Bradicardia: Un ritmo cardíaco lento (menos de 60 latidos por minuto)
- Arritmias supraventriculares: Ritmos anormales que se originan en las cámaras superiores del corazón (aurículas)
- Arritmias ventriculares: Arritmias que se originan en las cámaras inferiores del corazón (ventrículos)

- Bradirritmias: Ritmos lentos del corazón que generalmente están causados por una enfermedad en el sistema de conducción del corazón

c. *Factores fisiológicos que afectan la Frecuencia Cardíaca*

Los factores fisiológicos que afectan a la frecuencia cardíaca son todas aquellas variables fisiológicas o características propias del individuo que determinan su frecuencia cardíaca.

Para el desarrollo de esta investigación en particular, se ha contemplado el uso de las siguientes variables fisiológicas:

- Sexo: El ritmo cardiaco normal en el sexo femenino tiene entre 5 y 15 pulsaciones por minuto menos que del sexo masculino
- Edad: Conforme avanza la edad los latidos por minuto normales en una persona van disminuyendo.
- Peso y Altura: Las personas más altas tiene las pulsaciones más bajas que los más bajos y las personas de mayor peso tienen pulsaciones más bajas que las de menor peso.

IV. LEVANTAMIENTO DE DATOS

Esta investigación constará con datos reales otorgados por el departamento médico de la PUCESE, dichos datos incluirán variables tales como: sexo, edad, peso y altura, además se buscará saber si los miembros de esta población realizan actividad física, su intensidad y si consumen de tabaco.

La presente investigación está basada en el uso de un perceptrón usando el algoritmo BackPropagation, así que el tratamiento del error será con el error de cero para el tratamiento del valor de las variables a medir el y error cuadrático para la aplicación del back Propagation, [15].

V. METODOLOGIA

a. *Aplicación de una Red Neuronal de tipo Perceptrón para determinar si existe una frecuencia cardíaca elevada*

La primera aplicación de la red neuronal sobre el proyecto consiste en el entrenamiento de un Perceptrón para determinar en primera instancia si existe una frecuencia cardíaca elevada o fuera de lo normal, este procedimiento consiste en la base de las siguientes operaciones, puesto que uno de los factores más importantes a tomar en cuenta es la frecuencia cardíaca.

El perceptrón funcionará de manera que reciba los datos desde el dispositivo, sobre todo el pulso a cada instante, y al momento de detectar un pulso fuera de lo común se iniciará un proceso de recolección de los demás datos para enviárselos a la siguiente red neuronal.

Para las pruebas correspondientes se realizó una simulación mediante el programa MatLab, en el cual se codificó un Perceptrón simple y se lo entrenó con datos para definir su comportamiento.

b. *Aplicación del tipo de aprendizaje BackPropagation sobre una Red Neuronal Multicapa para el cálculo de un diagnóstico previo según las variables determinadas*

El proceso que continúa al punto anterior corresponde a una red neuronal multicapa entrenada mediante el tipo de aprendizaje BackPropagation, este procedimiento comienza luego de detectar una actividad cardíaca elevada o poco común para el usuario según los datos que se tengan del mismo, corresponde a una red neuronal capaz de determinar un diagnóstico sobre la situación actual de la persona para determinar su estado.

Para las pruebas realizadas se entrenó la red de manera que responda a varios datos como entradas incluyendo las variables mencionadas anteriormente y en base a esto y el pulso cardíaco calculado determine un diagnóstico del usuario.

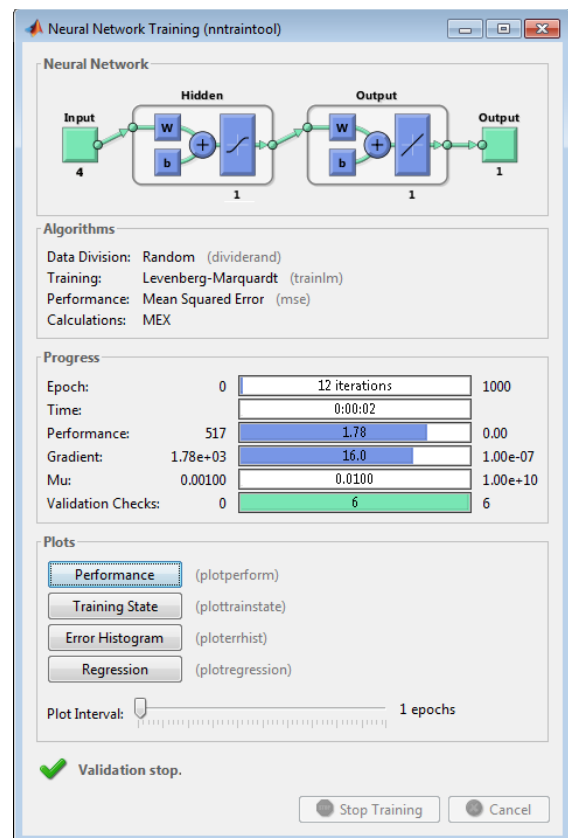


Fig. 5. Modeloy entrenamiento de Red Neuronal mediante la herramienta nntool MatLab

c. *Dispositivo Electrónico*

Para el prototipo y pruebas correspondientes del dispositivo se usa además varios sensores y dispositivos que al interactuar con el usuario permiten conocer en tiempo real los datos necesarios para los cálculos.

i. *Sensor de Pulso*

El sensor de pulso es un dispositivo plug-and-play para cualquier tipo de microcontrolador. Se optó por un modelo genérico y adaptable que funciona a manera de sensor de ritmo cardíaco óptico, una etapa de amplificación y un filtro

para el ruido, lo cual hace que su señal de salida sea confiable y estable. El consumo de corriente es bajo siendo de 4mA con una alimentación de 5V.



Fig. 6. Sensor de Pulso Modelo SEN_0386

ii. Módulo GPS SkyLab SKM53

El módulo SkyLab SKM53 fue diseñado para aplicaciones de sistemas de posicionamiento global GPS. El dispositivo está basado en el circuito MT3329 de MediaTek y cuenta con una antena GPS integrada al módulo.

El módulo SKM53 cuenta con 22 canales de rastreo y 66 canales de adquisición. Su precisión varía dependiendo del número de satélites con los cuales mantenga comunicación, siendo de 3 metros la especificación del fabricante.

Para su conexión con microcontroladores, el módulo SKM53 cuenta con un puerto serial asíncrono, con un conector de 6 pines, alimentación de 5 volts @ 50 ma. y señales con niveles TTL, utilizando el protocolo NMEA-0183, y operando a una velocidad default de 9600 bps.



Fig. 7. Módulo GPS SkyLab SKM53

VI. CONCLUSIONES

Mediante la presente investigación se concluye que es factible y viable la elaboración de un dispositivo capaz de medir el pulso cardíaco de una persona y en base a esto y varios factores asociados realizar un diagnóstico que determine el estado actual del usuario, y de ser una emergencia el dispositivo pueda ser de ayuda en la localización de puntos estratégicos como hospitales o farmacias para el posible tratamiento de la persona.

La importancia de las redes neuronales aplicadas a la medicina radica en la capacidad de entrenamiento que se puede tener, con los suficientes datos reales es posible obtener un resultado sumamente acertado, siendo de gran ayuda en casos de emergencia determinando diagnósticos y

posibles afecciones, que en este caso están relacionadas con el corazón.

VII. TRABAJOS FUTUROS

Esta investigación deja planteadas muchas posibilidades, desde la optimización y mejora del dispositivo en cuanto a hardware y software, hasta la implementación de varias características más en la medida de lo posible. El prototipo físico del dispositivo es un buen punto de partida, ya que a partir de esto sería posible determinar sus carencias y mejorarlas conforme se optimiza el código y se añaden funcionalidades.

Por otro lado, si se piensa en las redes neuronales, las aplicaciones posibles para estas son enormes, la investigación realizada tuvo buenos resultados incluso usando solo dos tipos de redes también se plantea como trabajo futuro para los autores la investigación más a fondo de este tema, ampliando las posibilidades de implementaciones futuras para el dispositivo.

REFERENCIAS

- [1] A. Barea, "Diseño, desarrollo y test de un prototipo de pulsera para adquirir el electrocardiograma y la onda de pulso," 2016.
- [2] A. Mariel, "Un electrocardiógrafo inteligente de bajo coste," 2014.
- [3] V. Pedro and V. Luis, "Monitor de ritmo cardíaco para dispositivos Android mediante un enlace Bluetooth," 2016.
- [4] E. Tintín, "Diseño y elaboración de un prototipo de monitor de signos vitales aplicando métodos no invasivos con comunicación de datos a dispositivos móviles," 2015.
- [5] E. Singer, "Un reloj de pulsera que controla la presión arterial," 2011. [Online]. Available: <https://www.technologyreview.es/s/2056/un-reloj-de-pulsera-que-controla-la-presion-arterial>. [Accessed: 27-Nov-2017].
- [6] L. Moisés, "Sistema Inteligente para la Toma de Pulsaciones Cardíacas de Forma Pasiva mediante Android Wear Intelligence System to Measurement Heart Rate with a passive method by means of Android Wear," 2015.
- [7] D. Matich, "Redes Neuronales: Conceptos Básicos y Aplicaciones.," 2001.
- [8] L. Flores, "Perceptrón," 2014. [Online]. Available: <http://www.tecnohobby.net/ppal/index.php/inteligencia-artificial/redes-neuronales/11-perceptron>. [Accessed: 27-Nov-2017].
- [9] J. McGonagle, G. Shaikouski, A. Hsu, J. Khim, and C. Williams, "Backpropagation," 2017. [Online]. Available: <https://brilliant.org/wiki/backpropagation/>. [Accessed: 27-Nov-2017].
- [10] Ma. Mazur, "Un ejemplo de Backpropagation paso a paso," 2015. [Online]. Available: <https://mattmazur.com/2015/03/17/a-step-by-step-backpropagation-example/>. [Accessed: 27-Nov-2017].
- [11] X. Basogain, "Redes Neuronales Artificiales Y Sus Aplicaciones," Med. Intensiva, vol. 29, no. 1, pp. 13–20, 2005.
- [12] "National Heart, Lung and Blood Institute," ¿Cómo se diagnostican las enfermedades del corazón?, 2012. [Online]. Available: <https://www.nhlbi.nih.gov/health-spanish/health-topics/temas/hdw/diagnosis>.
- [13] A. Mardal, "NEWS MEDICAL LIFE SCIENCES," ¿Cuál es Ritmo Cardíaco?, 2017. [Online]. Available: [https://www.news-medical.net/health/What-is-Heart-Rate-\(Spanish\).aspx](https://www.news-medical.net/health/What-is-Heart-Rate-(Spanish).aspx).
- [14] M. Robbin G. Cohen, "The Society of Thoracic Surgeons," Trastornos del Ritmo Cardíaco, 2016. [Online]. Available: <https://ctsurgerypatients.org/es/enfermedades-del-corazón-en-adultos/trastornos-del-ritmo-cardíaco>.
- [15] M. P. Gonzalez, "Error Cuadrático Medio de P redicción para M odelos Estructurales de Series Temporales," Estad. Española, vol. 34, no. 129, pp. 117–135, 1992.

Parámetros incidentes en el consumo eléctrico, referidos al sector residencial urbano

Juan Segura¹, FranyelitSuárez²

¹Universidad Tecnológica Indoamérica. ²Pontificia Universidad Católica del Ecuador-Sede Esmeraldas
 juanjoelsegura@yahoo.com, franyelit.suarez@pucese.edu.ec

Resumen— El presente trabajo de investigación presenta el análisis de los factores más resaltantes que condicionan el consumo eléctrico residencial urbano. En este estudio se muestran los parámetros cuantitativos condicionantes del consumo eléctrico. Este sector de análisis ha sido escogido debido a que existe información de manera desagregada de cuáles son los principales factores sociales y tecnológicos que determinan su comportamiento, crecimiento, con el objetivo de elaborar políticas en la gestión del consumo eléctrico. La demanda eléctrica considerada como la suma de las potencias de todos los equipos que se estén utilizando en cada uno de los instantes de un día completo, está relacionada con el consumo eléctrico, que no es sino el valor de potencia demandada por un consumidor determinado multiplicado por el tiempo en que dicha demanda se mantiene. En este informe se plantea el diseño de un modelo probabilístico de predicción del consumo eléctrico, tomando en cuenta principalmente los factores sociales y tecnológicos influyentes. El proceso estadístico de esta base de datos se realizó a través del programa de computación StatGraphics versión 4.1, por su amplia didáctica en la realización de cálculos y métodos asociados. Finalmente se realizó la correlación de las variables para clasificar los condicionantes de forma específica y lograr así determinar el consumo de las viviendas.

Palabras claves: consumo eléctrico, proceso estadístico, sector urbano, factores sociales.

I. INTRODUCCIÓN

La energía eléctrica se puede considerar como una de las materias prima, más utilizada en la sociedad moderna, se ha convertido en una necesidad de los pueblos, con el avance de nuevas tecnologías domésticas que demandan un mayor consumo de la misma. Generalmente cuando se adquiere un artículo electrodoméstico (televisor, refrigerador, lavadora, secadora, etc.), se está agregando un consumo adicional al diseño original del circuito eléctrico, lo cual puede generar un sobreconsumo o una sobrecarga que puede generar un peligro de incendio, producto del recalentamiento de los conductores y de todo el sistema de red eléctrica. Usualmente los riesgos eléctricos se producen por una excesiva conexión de equipos a una red que fue diseñada para un número moderado de los mismos. Benaventes et al [1] sugiere que la demanda eléctrica responde al precio que se cobra por la misma, y que por lo tanto los cortes de energía son innecesarios cuando hay un déficit en electricidad. En períodos normales se puede considerar que la demanda de energía depende casi solamente del ingreso de los consumidores y del precio relativo [2], [7]. En el siguiente trabajo de investigación se presenta el desarrollo de un análisis estadístico para la valoración del consumo eléctrico

residencial urbano en la provincia de Santiago de Cuba, Cuba. En este estudio se han considerado los parámetros sociales y tecnológicos que determinan el comportamiento del consumo eléctrico. Una vez realizadas las evaluaciones del sector urbano se propuso un modelo probabilístico de predicción del consumo, que permita tomar en consideración el crecimiento poblacional y las características sociales que puedan garantizar un buen servicio eléctrico de forma continua. En este trabajo se presenta un análisis estadístico de las condiciones del sector urbano, se propone un modelo matemático para la predicción del consumo y se evalúan las causas que afectan dicha situación.

II. DESARROLLO

1. Materiales y Métodos

Para el siguiente estudio se consideraron los principales factores sociales y tecnológicos que inciden en el consumo eléctrico urbano; tanto en su comportamiento, y crecimiento, con el objetivo de elaborar políticas en la gestión del mismo.

Para este trabajo se tomó como población objeto de estudio al sector sur de la ciudad de Santiago de Cuba. Cuba. El mismo se caracteriza por una alta densidad de población. En el mismo se pudo apreciar una serie de características, las cuales hablan a favor de una cierta equidad social en el conjunto residencial.

Fundamentos teóricos

Demanda Eléctrica: Se considera Demanda Eléctrica a la sumatoria de potencias consumidas por los equipos que se estén utilizando a cada instante del día y de la noche, esta magnitud se mide en Watt o múltiplos de ella.

Consumo Eléctrico: Se refiere al valor de potencia exigida por un consumidor multiplicado por el tiempo en que dicha potencia se mantiene, se expresa en Watt-horas o múltiplos de esta unidad.

Característica del Sector Residencial.

- 1- Posee un bajo factor de carga, pues durante el día la demanda es muy baja la cual se incrementa bruscamente en el horario nocturno, coincidiendo con el horario de máxima demanda.
- 2- El factor de carga depende del nivel de electrificación del país.
- 3- Posee altas tasas de crecimiento anual.

- 4- Sus consumos fluctúan con el clima y nivel de vida de sus habitantes.

Selección de factores a estudiar de forma cuantitativa

1. Mes (factores estacional)
2. Potencia instalada (factor tecnológico)
3. Ingreso total familiar (factor social-económico)
4. Número de personas/por viviendas (factor social-demográfico)

Selección de factores a estudiar de forma cualitativa

- 1- Ocupación
- 2- Nivel de escolaridad
- 3- Edad

Selección de la población objeto de estudio

La población objeto de estudio, la constituye el complejo urbano Sierra Maestra, ubicado en la zona sur de la ciudad de Santiago de Cuba, constituida por 5 edificios de dieciocho plantas cada uno, con un total de 526 departamentos. El mismo posee una alta densidad de población y una composición social heterogénea.

Cálculo del tamaño de la muestra en la población seleccionada

El tamaño de la muestra se determina, a través de la ecuación 1, arrojando un valor de n=100 departamentos

$$n = \left(\frac{Z * S}{E} \right)^2 \quad (1)$$

donde: n: tamaño de la muestra.

Z: variable estandarizada, para $\alpha = 0,05$; $Z = 1,96$

S: desviación estándar de la medición en el objeto de medición (metro contador) $S = 0,05$

E: error permisible tomado para el estudio. $E = 0,01$

Se considerarán variables dependientes el consumo eléctrico expresado en KW-h, y como variable independiente las definidas de manera cualitativa y cuantitativa, relacionadas con los aspectos sociales y tecnológicos.

Los valores correspondientes a la variable Consumo Eléctrico pertenecen al año de estudio 2014, esta información fue suministrada por la sucursal Vista Alegre, perteneciente a la Empresa Eléctrica de Santiago de Cuba, previo conocimiento del control, ruta y folio de cada consumidor seleccionado. Dicha selección fue realizada mediante la aplicación de un programa estadístico, donde se realizó un muestreo simple sin reemplazo en la selección de las viviendas objeto de estudio. Una vez seleccionadas las mismas, se procedió a evaluar las características de cada una de ellas, a través de un sistema de encuestas y de un levantamiento de carga instalada en las mismas. Se utilizó el software de computación StatGraphics versión 4.1 para el

procesamiento de datos, ya que constituye una herramienta muy útil para procesamiento de datos estadísticos con alta confiabilidad de respuesta. El conjunto de datos se aprecia en la tabla 1, donde se muestran los 12 modelos correspondientes a cada uno de los meses del año.

Tabla 1. Resumen de los Modelos Mensuales

Mes	R ² Ajustada (%)	Modelo mensual (kW-h)	Error estándar del estimado (kW-h)
Enero	91,1023	21,3712Pinst + 23,5399 No. Pers.	44,4967
Febrero	90,0709	16,5691Pinst + 20,3878 No. Pers.	39,8578
Marzo	91,6072	22,3367Pinst + 21,6926 No. Pers.	41,9842
Abril	90,1458	19,5692Pinst + 24,8360 No. Pers.	47,9193
Mayo	89,9815	24,2341Pinst + 20,2330 No. Pers.	46,3811
Junio	92,4096	27,8018Pinst + 22,4906 No. Pers.	44,5491
Julio	93,0441	26,2801Pinst + 22,8229 No. Pers.	41,6399
Agos.	93,4384	32,3750Pinst + 21,0059 No. Pers.	41,6792
Sep.	93,8472	26,2612Pinst + 23,6899 No. Pers.	39,5664
Oct.	93,5358	27,6168Pinst + 20,3710 No. Pers.	39,1861
Nov.	91,7972	25,7062Pinst + 21,0031 No. Pers.	43,6443

Dic.	90,8683	20,2742Pinst + 21,3052 No. Pers.	42,4683
------	---------	-------------------------------------	---------

1. Estudio de Correlación.

Se realizó un estudio de correlación entre el Consumo, como variable dependiente y las variables independientes de carácter cuantitativo; Mes, Potencia instalada, Número de personas/por viviendas e Ingresos. Se pudo comprobar que la variable Ingresos no correlaciona con las siguientes variables: Consumo, Mes y Potencia instalada. Ya que esta variable en el contexto de Cuba, posee connotaciones especiales, pues en el país está establecida una doble moneda circulante, por lo que el verdadero valor de Ingreso por cada vivienda no es posible de precisar, a través de las encuestas realizadas a las mismas.

verdadero valor de Ingreso por cada vivienda no es posible de precisar, a través de las encuestas realizadas a las mismas.

1. Análisis de Regresión Lineal Múltiple

Para la realización del estudio de Regresión Lineal Múltiple se propone la confección de 12 modelos, uno mensual y de esta forma cada uno de ellos tendrá implícitos la variabilidad estacional y mensual respecto al Consumo.

Datos empleados en la confección de los modelos.

- 1.-Consumo para cada mes específico de cada vivienda seleccionada.
- 2.-Potencia instalada.
- 3.-Número de personas/por vivienda.

Tabla 2. Resumen de los modelos mensuales.

Mes	R ² Ajustada (%)	Modelo mensual (kW-h)	Error estándar estimado (kW-h)
Enero	91,1023	C _{Ene} = 21,3712Pinst + 23,5399 No. Pers.	44,4967
Febrero	90,0709	C _{Feb} = 16,5691Pinst + 20,3878 No. Pers.	39,8578
Marzo	91,6072	C _{Mar} = 22,3367Pinst + 21,6926 No. Pers.	41,9842
Abril	90,1458	C _{Abr} = 19,5692Pinst + 24,8360 No. Pers.	47,9193
Mayo	89,9815	C _{May} = 24,2341Pinst + 20,2330 No. Pers.	46,3811

Junio	92,4096	C _J = 27,8018Pinst + 22,4906 No. Pers.	44,5491
Julio	93,0441	C _J = 26,2801Pinst + 22,8229 No. Pers.	41,6399
Agos.	93,4384	C _{Agt} = 32,3750Pinst + 21,0059 No. Pers.	41,6792
Sep.	93,8472	C _{Sep} = 26,2612Pinst + 23,6899 No. Pers.	39,5664
Oct.	93,5358	C _{Oct} = 27,6168Pinst + 20,3710 No. Pers.	39,1861
Nov.	91,7972	C _{Nov} = 25,7062Pinst + 21,0031 No. Pers.	43,6443
Dic.	90,8683	C _{Dic} = 20,2742Pinst + 21,3052 No. Pers.	42,4683

2. Validación de los modelos obtenidos a través del estudio de Regresión Lineal Múltiple.

La validación del modelo se debe realizar en elementos de la población seleccionada que cumplan los siguientes requisitos:

- 1.-Las viviendas objeto de validación se deben escoger al azar.
- 2.-Las viviendas objeto de validación no pueden formar parte de la muestra.
- 3.-Las viviendas objeto de validación deben poseer un consumo histórico mensual durante el año 2015 entre (50-250) kW-h.

Se validó el modelo comparando el valor real de consumo obtenido, a través de la facturación del mes en estudio, con el valor de consumo arrojado por la ecuación del modelo de Regresión para el mes analizado junto con su intervalo de predicción para observaciones futuras. Si el valor real de consumo está dentro del intervalo calculado el modelo es válido para el nivel de confianza seleccionado, un 95%.

El intervalo de predicción se obtiene a través de la ecuación (2)

$$\hat{Y}(X_0) \pm t \frac{\alpha}{2, n-p} \sqrt{\hat{\sigma}^2 [1 + X_0' (X'X)^{-1} X_0]} \quad (2)$$

donde:

\hat{Y} - Valor obtenido, a través del modelo.

$\hat{\sigma}$ - Desviación estándar de los residuos de los modelos, se emplea como límite de predicción para nuevas observaciones.

X_0' - La matriz transpuesta del vector columna correspondiente a los nuevos valores o valores futuros de las variable regresoras, en este caso: potencia instalada y número de personas.

$$X_0' = [1 \ X_{01} \ X_{02}]$$

X_0 - Vector columna de los valores futuros

$$X_0 = \begin{bmatrix} 1 \\ X_{01} \\ X_{02} \end{bmatrix} \quad (3)$$

$[X' X]^{-1}$ - Es la matriz inversa del producto de la matriz transpuesta por ella misma.

t - es la t de Student

α - **0,05** para un intervalo de confianza de un **95 %**

$n - p$ - Grados de libertad

n - Número total de observaciones

p - Parámetros del modelo.

Si se simplifica la ecuación 2 despreciándose el término $[1 + X_0'(X'X)^{-1}X_0]$ el cual forma parte de la cantidad subradical, traería como consecuencia un resultado menor considerándose el menor de los casos para el intervalo de predicción futura, la ecuación 4 muestra la simplificación obtenida:

$$\hat{Y}(X_0) \pm t \frac{\alpha}{2, n - p} \hat{\sigma} \quad (4)$$

Se propone para la validación del modelo tomar un 5% de la población seleccionada que cumpla con los requisitos citados anteriormente. El valor del 5% se ha tomado por tener un nivel de confianza de un 95% que permite lograr una respuesta con un error de hasta un 5%. La población seleccionada está constituida por 526 apartamentos, para lo cual el 5% lo constituyen 26 departamentos, con el objetivo de verificar si el valor real obtenido se encuentra dentro del rango para cada modelo mensual.

3. Estudio y análisis del Consumo Eléctrico respecto a las variables cualitativas

Para el cumplimiento del mismo se plantea un análisis de frecuencia de la variable ocupación, el cual tiene implícito las otras las otras dos variables respecto al consumo.

Análisis de Frecuencia para la Variable Ocupación.

Para la realización del presente estudio se establecen diferentes grupos de clases en dependencia de la ocupación que posean las personas integrantes de los núcleos familiares en las viviendas objeto de estudio y análisis.

Los grupos de clases propuestos son los siguientes:

- *Grupo A.* Está integrado por el conjunto de departamentos pertenecientes a la muestra de 100 viviendas donde la ocupación de sus habitantes se denomina trabajadores solos, o éstos asociados a estudiantes.
- *Grupo B.* Está integrado por el conjunto de departamentos pertenecientes a la muestra de 100 viviendas donde la ocupación de sus habitantes se denomina trabajadores, jubilados, estudiantes, amas de casa, niños, es decir, mixtos.
- *Grupo C.* Está integrado por el conjunto de departamentos pertenecientes a la muestra de 100 viviendas donde la ocupación de sus habitantes se denomina jubilados solos, o éstos asociados a estudiantes.

Para la muestra estudiada la composición de la misma de acuerdo a la ocupación de las personas es presentada en la tabla 3. Donde se observa un valor importante en la cantidad de trabajadores en comparación con la cantidad de amas de casa.

Tabla 3. Composición de la muestra

Ocupación	Nº de personas
• Trabajadores	167
• Estudiantes	63
• Jubilados	51
• Ama de casa	16
• Niños y adolescentes	22
• Total de la muestra	319

Determinación de datos estadísticos tanto para la muestra en su conjunto como para los diferentes grupos de clases.

Se determinaron los valores estadísticos de la muestra, expuestos en la tabla 4, adicional se muestra la mediana, el mínimo y el máximo del conjunto de datos

Tabla 4. Resumen de los valores estadísticos

Estadísticos	P/Muestra (kW-h)	Grupo A (kW-h)	Grupo B (kW-h)	Grupo C (kW-h)
Casos válidos	100	57	31	12
Media	1801,07	1673,11	2083,03	1680,5
Mediana	1769	1710	1996	1413
Mínimo	130	130	792	817
Máximo	5620	3020	5620	3151
Total	180107	95367	64574	20166

Valoración acerca del análisis de frecuencia para las variables cualitativas.

Los consumidores pertenecientes al Grupo B (grupo mixto), poseen las siguientes características:

- Tienen el mayor consumo promedio anual y mensual por viviendas respecto a la muestra.
- El grupo tiene el segundo lugar en el consumo con respecto al total de la muestra seleccionada.

Al comparar el consumo promedio anual del Grupo B en relación a los Grupos A y C, se observa que el primero arroja valores superiores. De acuerdo a los resultados anteriores se infiere que las viviendas pertenecientes a este Grupo B tienen mejores características respecto a los otros dos grupos que conforman la muestra, motivado a que van a poseer un mayor número de personas en sus casas, diversidad en las ocupaciones de sus integrantes, expresándose esto en un mayor consumo per cápita. De esta forma se corrobora la correlación existente entre consumo y número de personas/vivienda.

Los consumidores pertenecientes al Grupo C (Jubilados solos o asociados a estudiantes) poseen un consumo promedio mensual por vivienda y consumo promedio anual por vivienda prácticamente igual al Grupo A (Trabajadores solos o asociados a Estudiantes)

Por tal motivo se infiere la importancia de conocer la ocupación de las personas con vista a la realización de estudios encaminados al pronóstico del consumo y la demanda, pues los jubilados al tener una mayor permanencia en el hogar, contribuyen al incremento de ambas, independientemente de que ellos como grupo respecto a la muestra sean los de menor consumo.

III. RESULTADOS

Del estudio de Regresión Lineal Múltiple se obtuvieron los siguientes resultados:

- Se empleó un modelo lineal sin constante numérica como primer término, debido a que en ausencia de todas las variables independientes el consumo sería **0**, pues no existe en las viviendas una carga fija consumidora desvinculada de las variables independientes.
- Los modelos obtenidos van a poseer una **R²** ajustada en el rango (**89,9815– 93,8472**) % indicando que los mismos explican en ese rango la variabilidad del consumo en dependencia del mes que se estudie.
- Las variables que conforman el modelo están muy interrelacionadas, por tanto: el modelo se presenta como un todo, pudiéndose emplear en predecir el consumo, sin embargo, no puede ser usado para evaluar rigurosamente el aporte que tiene cada variable por separado en la respuesta total.

- El **P-value** en cada modelo toma una magnitud de **0,0000** indicando una alta significación estadística para un **95 %** de nivel de confianza.
- Los modelos obtenidos, a pesar de poseer una **R²** ajustada en el rango anteriormente citado (excelentes valores), sin embargo, tienen un Error Estándar del Estimado en el rango de (**39,1861–47,9193**) kW-h, los cuales se consideran elevados con vistas a la predicción de nuevas observaciones.

IV. CONCLUSIONES

1.-Se establece una correlación estadísticamente significativa entre los factores social y tecnológico, respecto al consumo total anual, excepto la variable social cuantitativa ingresos.

2.-Se elaboró un modelo matemático correspondiente a cada mes del año para realizar una predicción probabilística acerca del consumo en una vivienda del sector residencial urbano, determinando el nivel de influencia de los parámetros cuantitativos seleccionados de manera conjunta y no individualmente.

3.-Se empleó un modelo lineal sin constante numérica como primer término, debido a que en ausencia de todas las variables independientes el consumo sería cero.

4.-Los modelos obtenidos van a poseer una R cuadrada ajustada en el rango (89,9815-93,8472) % indicando que los mismos explican en ese rango la variabilidad del consumo en dependencia del mes que se estudie.

5.-El modelo a pesar de resultar insuficiente para realizar en general pronósticos del consumo en el sector residencial urbano por carecer de otros elementos fue validado en 26 viviendas de la población seleccionada, representando 312 meses, de los cuales en sólo 10 casos el valor real no pertenecía al intervalo de predicciones futuras.

REFERENCIAS

- [1] Benavente, J., Galetovic, A., Sanhueza, R. y Serra, P. *Estimando la demanda por electricidad en Chile: el consumo es sensible al precio.* (2005) Vol. 42. pp. 31-61.
- [2] Marshall, D. *El consumo eléctrico residencial en Chile en 2008.* (2010) Vol. 47. pp. 57-89.
- [3] Puentes, V. *Demanda eléctrica.* Conferencia Nacional de Demanda Eléctrica, La Habana, 2002.
- [4] Martínez Serna P. Trabajo de Grado. Proyección de La Demanda. Antecedentes. Necesidad e Importancia. Universidad Nacional de Colombia. Sede Manizales, 2003.
- [5] De Barros Neto, Benicio; Spacino Scarminio, Leda; Edward Bruns, Ray. *Como Fazer Experiments.* II Ed. página 58. Editora Unican 2002.
- [6] Fierro, G., Serra, P. *Un modelo de estimación del costo de falla eléctrica: caso Chile.* Cuaderno de economía. Año 30, No. 90 (Agosto 1993), pp. 247-259
- [7] IDEA. Secretaría general departamento de planificación y estudio. (2011). *Análisis del consumo energético del sector residencial en España.*
- [8] Agostini, C., Pottier, M., Saavedra, E. (2009) *LA demanda residencial por energía eléctrica en Chile.* CASEN.

La vinculación entre la universidad y el sector productivo industrial

Martínez Danilo
pcmarket1@hotmail.com,
Corporación Nacional de Electricidad, CNEL

Resumen — En este trabajo se fundamenta la conceptualización de la vinculación de la educación superior con respecto a los sectores productivos industriales para relacionarlos con los procesos académicos que deben ser satisfactorios en una sociedad cuando cubre la necesidad de educar a la ciudadanía, sin embargo, el tema de globalización y tecnología está encaminada a la producción y comercialización de productos y servicios para cubrir esa necesidad de mejorar la calidad de vida. Sin embargo, los datos estadísticos aquí presentados sobre la oferta académica existen, 30 universidades públicas, 26 universidades privadas, 85 Instituciones técnicas y tecnológicas públicas, 89 instituciones técnicas privadas y entre ellas hay 2.804 carreras para la elección.

El sector productivo transforma lo técnico con lo material y en ese sistema está sujeto a modificaciones e innovaciones que ejerce la producción de un bien o servicio, de aquí nace la pregunta ¿Cuál es la vinculación entre la universidad y el sector productivo industrial? Por lo tanto, es necesario en el presente estudio es identificar el significado de vinculación, establecer la importancia de la vinculación académica. Además, determinar la relación de la vinculación académica con el sector productivo. Se concluye afirmando que la vinculación puede convertirse en una herramienta eficaz de promoción académica de la Universidad de acuerdo a las necesidades del sector industrial.

Palabras clave: Vinculación, Universidad, Sector Productivo, Industrial.

I. INTRODUCCIÓN

El significado de vincular es unir, atar, articular, enlazar esfuerzos, trabajos y objetivos, por lo tanto, la vinculación es: “el proceso integral que articula las funciones académicas de la docencia, investigación y extensión de la cultura, así como los servicios de las instituciones de educación superior, para su interacción eficaz y eficiente con el entorno socioeconómico” [1]

Es importante la vinculación académica porque es una forma de apoyar al desarrollo de mayores lazos entre la universidad y el sector productivo con los proyectos de fortalecimiento a la innovación tecnológica en el sector productivo, entre ellos están:

- Reconocer que la vinculación permita la convergencia de esfuerzos y voluntades en la búsqueda de soluciones a la compleja problemática que enfrenta el país.
- Mostrar su disposición a participar en proyectos conjuntos de mutuo beneficio.
- Señalar la necesidad de establecer mecanismos de colaboración permanente.
- Considerar necesario buscar alternativas de financiamiento gubernamental para los proyectos de investigación y desarrollo tecnológico que benefician a los micros y pequeñas empresas.

- Destacar que la vinculación es una oportunidad de complementación entre las empresas y las instituciones educativas en la introducción de mejoras en los procesos productivos y en el desarrollo de productos, en apoyo a la consolidación del aprendizaje empresarial, y

- Admitir la necesidad de eliminar la desconfianza derivada del desconocimiento mutuo, de las diferentes lógicas de trabajo y de las diferencias en misiones y fines institucionales

En este artículo se explica la metodología utilizada, el diseño de la plataforma, los resultados obtenidos de información secundaria dispuestas por el Instituto Nacional de Estadísticas y Censos, finalmente se presentan las conclusiones.

La vinculación debe estar definidas por los objetivos estratégicos y acciones institucionales que involucran al personal académico, esta relación es de intercambio y cooperación entre las instituciones públicas y privadas, en este caso la universidad debe promover el avance académico, el desarrollo científico y tecnológico y la solución de problemas concretos planteados en convenios y programas con el sector productivo. En estos últimos 20 años se han producido cambios entre las Instituciones de Educación Superior con la sociedad, porque orientaban actividades con la información que provenía de su interior, generando un considerable aislamiento con el fin de proteger la autonomía universitaria según la Asociación Nacional de Universidades e Instituciones de Educación Superior de México. [2]

Las acciones más relevantes que en conjunto integran el ámbito de la vinculación corresponden a la impartición de cursos de educación continua presenciales y a distancia, la realización de reuniones académicas y eventos conjuntos, los servicios de asesoría y consultoría, la formación y capacitación de trabajadores, la realización de proyectos de investigación para el desarrollo de procesos y productos y el diseño de programas de formación de investigadores para resolver problemas específicos del sector productivo.

Es necesario proponer nuevas políticas en el sector educativo superior que por un lado permitan mejora la vinculación con el sector productivo, de tal manera que provoquen la liberación de las capacidades potenciales de los estudiantes profesores e investigadores y que por otro lado aporten ingresos adicionales tanto a sus miembros como a las instituciones educativas de educación superior.

II. DESARROLLO

2. Metodología

Para lograr los objetivos propuestos se realizó la metodología que consistió en un estudio documental y bibliográfico. El primero se orientó a conformar un marco de referencia que

facilitara la delimitación precisa del problema, permitiera analizar lo que se sabía sobre el tema, definir los conceptos clave involucrados en la investigación e identificar las posibles variables involucradas en el tratamiento del problema. Esto permitió generar una perspectiva de análisis que fue fundamental para elaborar el instrumento y analizar la información secundaria obtenida. Se apoyó información con el Instituto Nacional de Estadísticas y Censos, 2016. [3]

III.DISEÑO

En prospectiva, como abrir el conocimiento a la realidad nacional los responsables de la vinculación en las macro-universidades le apuestan a la investigación como el factor primordial para el desarrollo y consolidación de dichos enlaces con el sector privado. Esto, porque la generación y aplicación de conocimientos, parte fundamental de la misión educativa, tiene su origen primigenio precisamente en la búsqueda de respuestas, en la aportación de explicaciones, en la solución de los problemas más apremiantes del entorno. Sin embargo, para participar con garantía de éxito en la economía del siglo XXI, es importante entender que se está ante una realidad en la cual los competidores están presentes en todas partes, tanto a nivel local como global, por tal razón los empresarios deben prestar mucha atención a los diversos factores y aspectos que inciden en el mejoramiento o sostenibilidad de la competitividad. [4]

Pero esta realidad responde a los intereses de los universitarios el querer vincularse con una empresa: conseguir dinero, mejor entrenamiento, evitar tiempos en trámites de

selección, resolver problemas, acceder a la infraestructura industrial y acceder a fondos que apoyen la vinculación.

Así mismo existen intereses de los empresarios industriales por la vinculación con las universidades pueden ser: acceder a mano de obra calificada, conseguir asesoría de la ciencia actual, contar con soporte técnico, acceder a las instalaciones universitarias, aumentar prestigio e imagen, ser considerados empresas amigables, usar recursos baratos y económicos, resolver problemas específico.

La oferta de cupos por carrera, sede, modalidad, jornada y régimen (semestral o anual) consiste en el número de estudiantes que cada institución de educación superior está en condiciones de admitir en el primer nivel de carrera para cada periodo académico. La institución de educación superior determinará el número mínimo de estudiantes por curso para su apertura en el periodo académico correspondiente. [5]

El diagrama de oferta de instituciones de educación superior en el Ecuador, está en la Fig. N°1.



Fig. N° 1 Oferta de Instituciones de Educación Superior

Fuente: <http://www.forosecuador.ec/forum/ecuador/educación-y-ciencia>

Para la oferta académica existen, 30 universidades públicas, 26 universidades privadas, 85 Instituciones técnicas y tecnológicas públicas, 89 instituciones técnicas privadas y entre ellas hay 2.804 carreras para la elección. De acuerdo a los reportes de resultados del Sistema Nacional de Nivelación y

Admisión – SNNA se identificó que existe una concentración de demanda en 20 carreras, estas 20 carreras representan el 52% del total de demanda de carreras en las Universidades Públicas.

La carrera con mayor demanda en Universidades Públicas del Ecuador es enfermería según la figura N° 2 siguiente:

20 Carreras más demandadas en Universidades Públicas Julio 2016

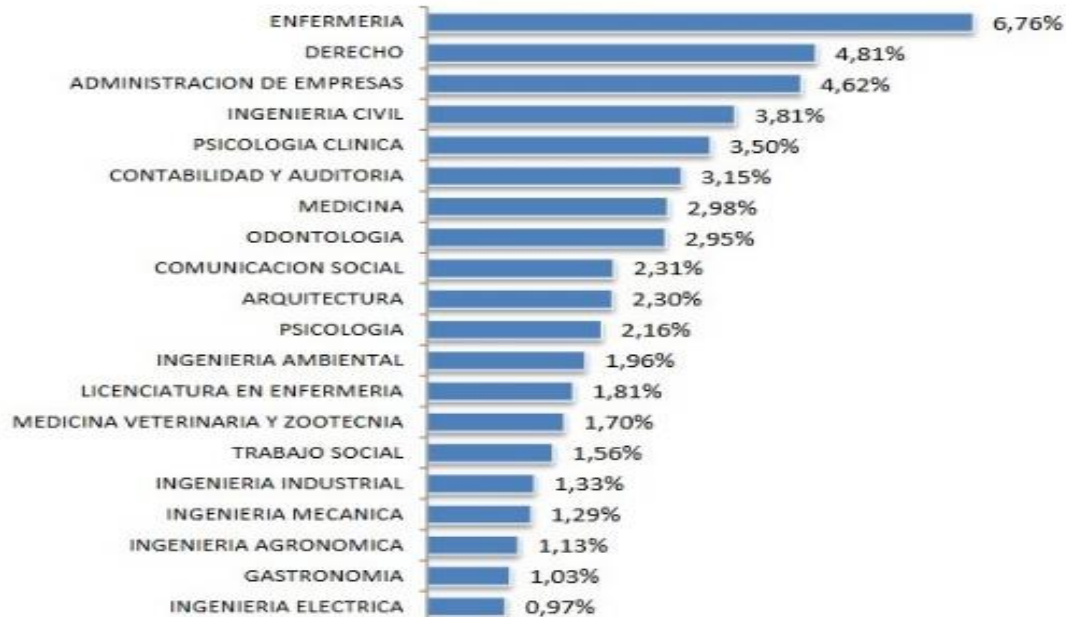


Fig. N° 2 Las 20 carreras más demandadas en Universidades Públicas.

Fuente: Base de datos del SSNA

El reporte del proceso ENES de marzo y septiembre del año pasado, de la Secretaría de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) por medio del sistema Nacional de Nivelación y Admisión (SNNA), arroja como resultado que las carreras con mayor demanda en los aspirantes para ingresar a la universidad, institutos técnicos y tecnológicos públicos, son las siguientes:

Carreras más solicitadas por los aspirantes que desean ingresar a las universidades públicas

Las carreras universitarias que vamos a nombrar a continuación, gozan de un afecto especial en los aspirantes y futuros universitarios.

- Enfermería
- Ingeniería civil
- Derecho
- Comunicación social
- Administración de empresas
- Psicología clínica
- Medicina
- Arquitectura
- Odontología
- Ingeniera ambiental
- Contabilidad y auditoría
- Economía
- Licenciatura de enfermería
- Ingeniería comercial
- Ingeniería industrial
- Psicología
- Medicina veterinaria y zootecnia
- Ingeniería mecánica
- Ingeniería agronómica
- Trabajo social

Las carreras más solicitadas por los aspirantes que desean ingresar a institutos técnicos y tecnológicos

Las carreras que nombraremos a continuación muestran claramente el gusto de los postulantes en cuanto a sus inclinaciones en estas carreras y futuras profesiones al culminar esta etapa de educación superior.

- Análisis de sistemas
- Hotelería y turismo
- Mecánica automotriz
- Diseño de modas
- Electromecánica
- Informática mención análisis de sistemas
- Mecánica industrial
- Agropecuaria
- Electrónica
- Agroindustria
- Diseño gráfico
- Electricidad
- Electricidad industrial mención electricidad de potencia
- Administración mención contabilidad y auditoría
- Electricidad industrial, mención en electromecánica
- Administración de empresas mención contabilidad
- Informática mención mantenimiento e instalación de redes
- Informática mención ensamblaje de equipos de computación
- Informática análisis de sistemas

- Asistente de ingeniería civil

En universidades públicas o en institutos técnicos y tecnológicos, lo importante es darles continuidad a los estudios, y de esta manera ir avanzando en cada etapa, adquiriendo los conocimientos y herramientas que se volverán cómplices en nuestra vida, una complicidad que te auguran un futuro prometedor y lleno de saberes. [6]

IV.RESULTADOS EXPERIMENTALES

A continuación, se presenta la clasificación de las empresas que se da por volúmenes de ventas anuales y la cantidad de personal que laboran.

CLASIFICACIÓN DE LAS EMPRESAS-VOLÚMENES DE VENTAS ANUALES

Clasificación de las empresas	Volúmenes de ventas anuales	Personal ocupado
Micro empresa	Menor o igual a 100.000	1 A 9
Pequeña empresa	De 100.001 a 1'000.000	10 A 49
Mediana empresa "A"	De 1'000.001 a 2'000.000	50 A 99
Mediana empresa "B"	De 2'000.001 a 5'000.000	100 A 199
Grande empresa	De 5'000.001 en adelante	200 en adelante

Nota: Estratos de Ventas, según la Comunidad Andina de Naciones. Decisión 702, Artículo 3.

De las consideraciones expuestas es importante tomar en cuenta la evaluación que ha tenido el número de empresas partiendo del año 2012 con un total de 749.912 para crecer en el año 2016 es decir después de 5 años 11,12 % llegando a un total de 843.745 empresas.

EVOLUCIÓN DEL NÚMERO DE EMPRESAS EN EL PERÍODO 2012 - 2016

Año	Nro. Empresas
2012	749.912
2013	826.804
2014	860.394
2015	858.835
2016	843.745

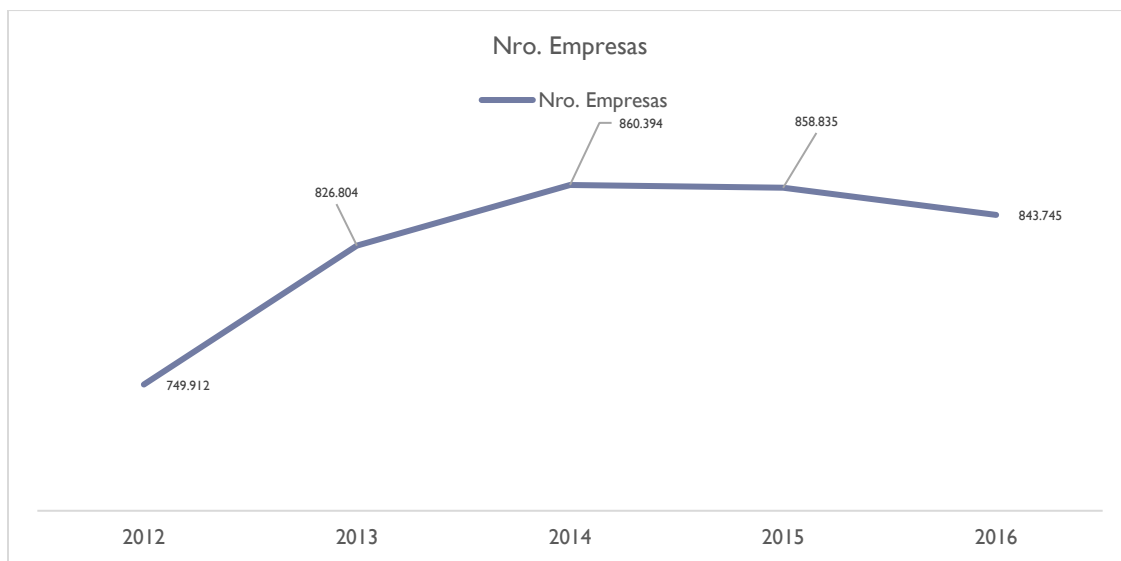


Fig. N° 3 Evolución del número de empresas en el período 2012 - 2016

Según se observa el comportamiento del indicador en la figura N° 3 de la evolución de las empresas en el período 2012-2016. El total de empresas del universo 843745 DICE, comprende todas las unidades económicas que registraron ventas en el SRI y/o registraron personal afiliado en el IESS o perteneciendo al RISE pagaron impuestos sobre sus ingresos en el SRI, en el año 2016.

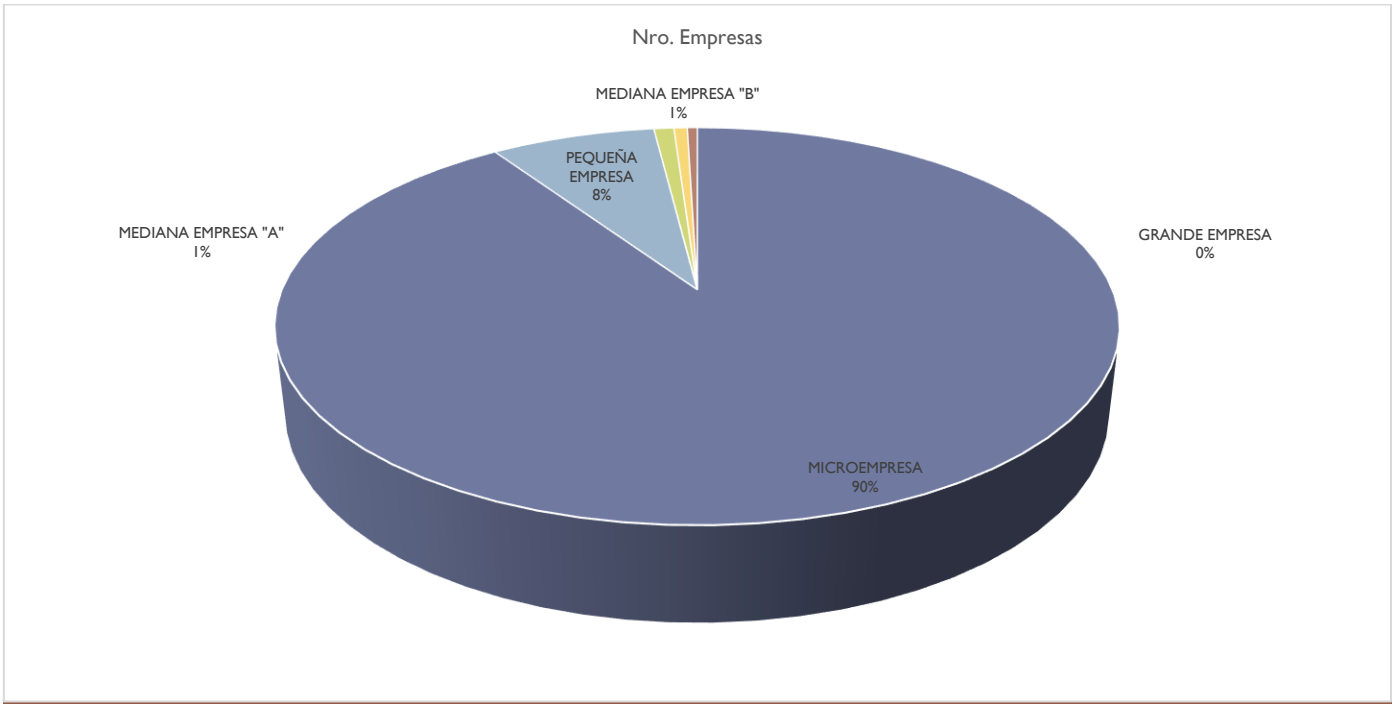


Fig. N° 4 Numero de empresas por el tipo

En la figura N° 4, según (Instituto Nacional de Estadísticas y Censos, 2016) se observa el porcentaje que tiene cada empresa en el mercado.

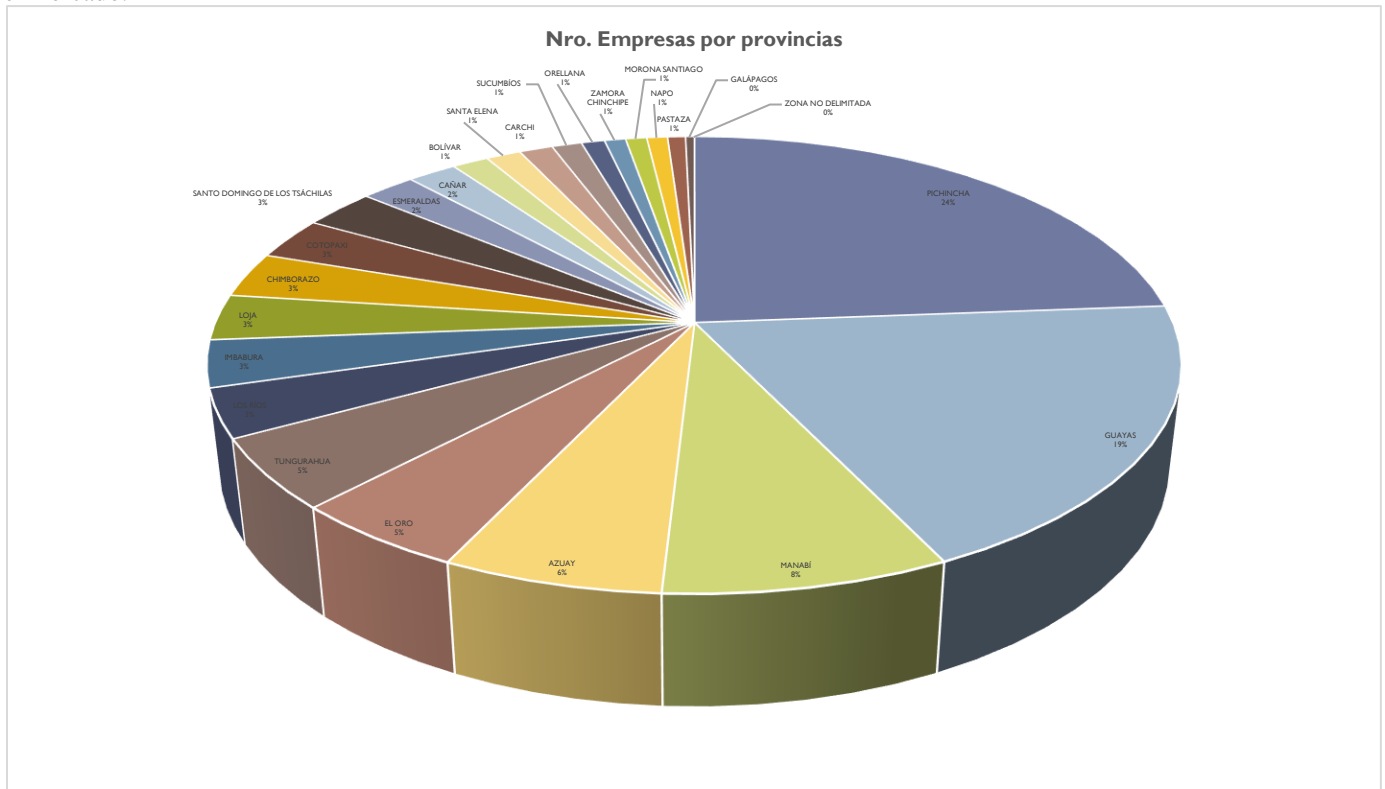


Fig. N° 5 Números de empresas por provincial.

En la figura N°5 indica el número de empresas por provincia donde se observa el porcentaje que existe en cada una de ellas de un total de 843745 de empresas.

Es significativo destacar que, libremente de las acciones que realicen las autoridades administrativas, la solides de la vinculación universidad y la relación con el sector productivo se da en términos de los intereses y gestiones individuales de los docentes con los gerentes de las empresas, estas gestiones a nivel mayor solo dan sus frutos en acuerdos de carácter general, difíciles de consolidar a menos que se acompañen de proyectos de factibilidades realizables en el corto y mediano plazo. [7]

Para justificar el presupuesto a las necesidades de la universidad se debe considerar las áreas o especialidades posibles para invertir en la formación académica dependiendo que la empresa requiere. Estas pueden ser las prioritarias: biotecnología, alimentos, química, farmacéutica, electricidad, electrónica, metal, mecánica, ecología y materiales, sin dejar la exportación e importación que es la base para generar ingresos y utilidades en las empresas.

Entonces se trata no únicamente de favorecer la realización de esta función sustantiva de las universidades –la investigación–, sino también de estrechar los lazos entre las instituciones de educación superior y las organizaciones lucrativas y no lucrativas para emprender actividades conjuntas, maximizar recursos, articular estrategias y acciones, intercambiar visiones y establecer consensos. [8]

En el área de tecnológica juega un papel importante la informática, lo digital aquí debe estimular la innovación y la creatividad de crear aplicaciones o software para la reducción en tiempos de producción siendo eficaces y eficientes en el uso de los recursos.

El universitario o prospecto de colaborador en una empresa debe egresar con una experiencia en las especialidades que necesiten las empresas en investigación y desarrollo para desempeñar eficazmente las funciones que se les asigne. Durante su formación debe practicar en empresas que estén vinculadas con la universidad he inclusive la universidad debe crear empresas universitarias donde los estudiantes practiquen en tiempo real y en sitio los problemas que tendrán que dar una solución.

IV. CONCLUSIONES

1. El graduado de una universidad tiene un conocimiento básico que puede ser complementado con especialización del área orientándolo profesional o académica, En esta oportunidad, no se trata únicamente de un reordenamiento de los contenidos entre grado y posgrado, es decir, de un traspaso de contenidos de las carreras de grado a las carreras de grado a las de posgrados, también significa en algunos casos, la eliminación de aquellas orientaciones, materias y temas que no demanda el mercado
2. En orden de promover la adopción de alianzas estratégicas para el ámbito de la vinculación universitaria, se considera necesario aportar también una serie de directrices más precisas que permitan a las universidades la ejecución apropiada de sus acciones
3. Fortalecer Programas de apoyo a la formación profesional de los estudiantes; Gestionar Programa de apoyo a la inserción laboral; Promover Programas emprendedores; Ejecutar Programas de investigación e innovación

tecnológica; Aplicar Programas de comunicación e información.

4. Con lo anterior indicado se deben establecer nuevas políticas educativas, definición de las necesidades de las empresas, inclusive formar consejos consultivos, dando paso a la creación en las instituciones redes interdisciplinarias.
5. Conseguir Contratación directa de trabajos a instituciones educativas. Algo similar también la contratación de estudiantes y pasantes en empresas, impulsando la participación de empresarios y gerentes en actividades académicas, generando apoyos de la Asociación de empresas consultoras a programas de vinculación que estén bien definidos y por ultimo un ofrecimiento mutuo de apoyo en áreas de coincidencia o sus similares

REFERENCIAS

1. Alcántar Enriquez VM. La vinculación como instrumento de imagen y posicionamiento de ls instituciones de educación superior. Redir ed. Mexicali, Baja California: Revista electrónica de Investigación Educativa; 2004.
2. Asociación Nacional de Universidades e Instituciones de Educación Superior. La Educación superior en el siglo XXI. Líneas estraégicas de desarrollo. México D.F.; 2010.
3. Instituto Nacional de Estadísticas y Censos. Censo poblacional. Censo. INEC, Departamento de Estadísticas; 2010.
4. Secretaria Permanente del SELA. Visión prospectiva de las Pequeñas y Medianas Empresas (PYMES) Respuestas ante un futuro complejo y competitivo Caribe SELyD, editor. Caracas-Venezuela: Sela; 2010.
5. Carrillo Hoyo E. Sobre la relación de las universidades con el sector productivo de bienes y servicios México D.F.: Dirección General de Investigación y Superación Académica DGICSA; 2010.
6. Haaz Mora H. La vinculación de las Instituciones de Educación Superior con los sectores sociales y productivos. México D.F.: UNAM; 2015.
7. Moreno Zagal M. Estrategias de vinculación de las universidades mexicanas con las empresas México D.F.: Congreso Nacional de Investigación Educativa; 2014.
8. Rikap C. La vinculación de la universidad con el sector productivo. Transferencia tecnologica Medellín-Colombia: ISSN 1657-4206 ; 2012.

Minería de opiniones de un personaje público en el Ecuador

Herrera Flores, Boris Enrique
bherrera@uce.edu.ec
Universidad Central del Ecuador

Resumen—La presente investigación es el estudio de las técnicas de minería de opiniones, enfocada a obtener información de un personaje público en el Ecuador, determinando indicios de polaridad a su gestión en forma positiva, negativa o neutra, resultado que le permitirá a dicho personaje público tomar decisiones sobre su accionar en función de una imagen de servicio a la comunidad. La extracción de las opiniones en redes sociales y técnicas basadas en Tecnologías del Lenguaje Humano posibilitó la interpretación de los datos polarizados precisando parámetros de relevancia a la opinión resultante enfocados a la toma de decisiones, procesamiento que se adapta a los nuevos formatos de comunicación logrando la interpretación y valoración de la opinión. Twitter fue la plataforma para la captura de textos por medio de un API, que luego del procesamiento del lenguaje natural se obtuvieron resultados de indicios de la popularidad del personaje.

Palabras clave: Procesamiento del lenguaje natural, minería de opiniones, análisis de sentimientos, tecnologías del lenguaje humano, clasificación de opiniones.

I. INTRODUCCIÓN

Actualmente la red social es la plataforma Web 2.0 para que los usuarios expresen sus opiniones, puntos de vista en forma inmediata con una participación frecuente usando un lenguaje informal, identifico la necesidad de procesar estos párrafos, oraciones y palabras precisando grados de relevancia a la informalidad al definir polaridad positiva, neutra o negativa, usando las Tecnologías del Lenguaje Humano.

Este estudio se enfoca a la comprensión de las expresiones subjetivas y el lenguaje informal usado por el Ecuatoriano al opinar sobre un personaje público a su accionar de servicio a la comunidad, donde las actuales tecnologías de lenguaje humano no son directamente aplicables y es necesario incorporar métodos y herramientas para el procesamiento del lenguaje natural por medio de la experimentación de diferentes técnicas y sistemas genéricos de análisis de sentimientos que extraigan primero la fuente de información para luego dar un tratamiento con métodos que determinen la polaridad de la opinión.

El objetivo es la mejora de los recursos, técnicas y herramientas que modelan el lenguaje subjetivo e informal que genera la red social Twitter, con el tratamiento del lenguaje emocional en un entorno de subsistemas inteligentes de procesamiento para la recuperación, tratamiento, comprensión y descubrimiento de la información apto para la toma de decisiones.

II. BASES CONCEPTUALES

Twitter, es una aplicación web gratuita de microblogging que recibe un estimado de 313 millones de visitas al mes, que reúne las ventajas de los blogs, las redes sociales y la mensajería instantánea, esta nueva forma de comunicación permite a sus usuarios estar en contacto en tiempo real con personas de su interés a través de mensajes breves de texto a los que se denominan Tweets, por medio de una sencilla pregunta: ¿Qué estás haciendo?. Los usuarios envían y reciben Tweets de otros usuarios a través de breves mensajes que no deben superar los 140 caracteres, vía web, teléfono móvil, mensajería instantánea o a través del correo electrónico; e incluso desde terceras aplicaciones.

Una API de Twitter (siglas en inglés “Application Programming Interface”), es un conjunto de reglas y especificaciones que las aplicaciones pueden seguir para comunicarse entre ellas, sirviendo de interfaz entre programas diferentes de la misma manera que la interfaz de usuario facilita la interacción humano-software. Twitter tiene tres tipos de APIs: REST API, Search API y Streaming API:

REST API, ofrece a los desarrolladores el acceso al core de los datos de Twitter. Todas las operaciones que se pueden hacer vía web son posibles realizarlas desde el API, soporta los formatos: xml, json, rss, atom.

Search API, suministra los tweets con una profundidad en el tiempo de 7 días que se ajustan a la consulta demandada. Es posible filtrar por, cliente utilizando, lenguaje y localización, no requiere autenticación y los tweets se obtienen en formato json o atom.

Streaming API, establece una conexión permanente por usuario con los servidores de Twitter y mediante una petición http se recibe un flujo continuo de tweets en formato json.

Con Search API y en el REST API existe una limitación de 150 peticiones a la hora por usuario o por IP si la llamada no estuvo autenticada.

El análisis de sentimientos, de textos en las redes sociales es el proceso que determina el tono emocional que hay detrás de una palabra determinada, si una frase contiene una opinión positiva, neutra, o negativa sobre un producto, marca, institución, organización, empresa, evento o persona, el objetivo es extraer aquellos términos semánticos que expresen un sentimiento en particular para conocer la opinión, las actitudes y las

expectativas sobre un tema en concreto así como para analizar el comportamiento de los usuarios ante algún mensaje y, por tanto, determinar su impacto o poder anticipar su reacción.

La determinación de la polaridad, consistente en determinar cuándo una opinión es positiva, negativa o neutra con respecto a la entidad a la cual se está refiriendo desde dos enfoques diferentes dependiendo del tipo de método que se utiliza, como son:

- Método de clasificación supervisada: Los rasgos extraídos del texto y el método de aprendizaje determinan cuando este pertenece a la clase positiva o negativa.
- Método de clasificación no supervisada (clustering): Tienen en cuenta la presencia de palabras con orientaciones conocidas que son obtenidas de diccionarios o corpus como, por ejemplo, las palabras excelente o alegre, que son representativas de expresiones polares positivas.

Existen varias técnicas al momento de estimar un texto y su posible polaridad positiva, negativa o neutra. Al utilizar software en la minería de opiniones se encuentran dos funciones: `classify_polarity` y `score_sentiment`

`classify_polarity`, Utiliza un marco probabilístico, tiene sus fundamentos en el Teorema de Bayes para calcular la

La metodología a utilizar es SEMMA creada por SAS Institute, que fue propuesta especialmente para trabajar con el software SAS Enterprise Miner. Si bien en la comunidad científica se conoce a SEMMA como una metodología, en el sitio de la empresa SAS se aclara que este no es el objetivo de la misma, sino más bien la propuesta de una organización lógica de las tareas más importantes del proceso de minería de datos. SEMMA establece un conjunto de cinco fases para llevar a cabo el proceso de minería:

Sample (Muestreo), Explore (Exploración), Modify (Modificación), Model (Modelado) y Assess (Evaluación).

Para el esquema general al detectar la polaridad de opiniones, se debe tomar en cuenta las palabras considerando su sentido correcto en un contexto determinado:

- Identificar Tokens: En esta primera etapa se realiza un pre- procesamiento del texto de las opiniones y devuelve los términos que aportan información útil, o definido como limpieza de los datos como por ejemplo duplicados generados por trolls
- Desambiguar lexicalmente cada token: En la segunda etapa se parte de cada término que aporta información útil, este le da forma, lematiza y desambiguan lexicalmente.
- Obtener todas las acepciones de cada palabra: Esta etapa parte del listado de términos lematizados y desambiguados y devuelve todas las acepciones del

probabilidad de una clase (positiva, negativa). Es una técnica de clasificación y predicción supervisada, entre sus ventajas está que su implementación es muy fácil y obtiene buenos resultados de clasificación en la mayoría de los casos:

$$p(C|F_1; \dots; F_n)$$

Donde C representa un valor positivo (+) o negativo (-) y F₁; F₂; ...; F_n factores que representa las palabras de un tweet, donde se busca establecer la probabilidad a priori de que un tweet sea positivo o negativo.

`score_sentiment`, es una función muy simple de predicción no supervisada que asigna una puntuación simplemente contando el número de ocurrencias de positivos y negativos en un tweet comparando con un diccionario de palabras conocido como corpus.

$$\text{score} = \text{sum}(\text{PalabrasPositivas}) - \text{sum}(\text{PalabrasNegativas})$$

El éxito de esta función es tener la mayor cantidad de palabras positivas y negativas, existen algunos diccionarios ya cargados con estas palabras uno de ellos es MPQA, LIWC.

III.METODOLOGÍA

término en el idioma en que se esté realizando el análisis.

- Clasificar cada token en positivo y negativo: Una vez obtenidas todas las acepciones de cada término, en esta etapa se propone determinar la polaridad del término de acuerdo a los algoritmos que se empleen, finalmente la polaridad de la opinión se determina por el número de palabras positivas y negativas que contiene.
- Evaluar la opinión: Si el número de palabras positivas que contiene una frase es mayor se considera la frase como positiva y en caso contrario es negativa.

Los índices de relevancia identificados en twitter son:

- Ratio seguidores/seguidos: Los usuarios que tienen un ratio de seguidores/seguidos cercano al 1 suelen contar con el efecto followback: muchas cuentas les siguen para ganar seguidores fácilmente. Para encontrar su relevancia usaremos una métrica basándonos en el ratio de seguidores/seguidos, pero esta será solo una parte, la parte fundamental para encontrar la relevancia se encuentra en el número de retweets que ha tenido el tweet asignándole un factor de multiplicidad de 6.5 al valor original, así mismo al campo favoritos que indica cuantas veces el tweet ha sido marcado como favorito, se le asignó un factor de multiplicidad de 3.5.
- Retweets/nº seguidores: Los retweets son un indicador universal de que lo que dices es interesante o digno de

atención. Cuantos más retweets tengas, más interesante o relevante es lo que dices o quien eres. El número de tweets es un factor importante, ya que no es lo mismo conseguir 100 retweets con un solo tweet que hacerlo a lo largo de un mes con 300.

El software para la extracción de datos utilizada es #TAGS, una aplicación que funciona con Google Docs, que permite entre otras cosas conocer los tweets y las

conversaciones que ha tenido un usuario determinado ofreciendo características adicionales del tweet como ubicación geográfica, id usuario, seguidores, favorito, retweets, etc.

Del estudio realizado aplicando la metodología indicada anteriormente, el personaje público obtuvo 71.845 tweets en periodo del 27 de diciembre del 2016 al 15 de febrero del 2017.

IV.RESULTADOS EXPERIMENTALES

Del estudio realizado de 71.845 tweets usando el software de JMP de SAS se aplicó la limpieza a los datos, eliminando links en el texto, tweets vacios, duplicados aplicando la fórmula de la Figura N°1

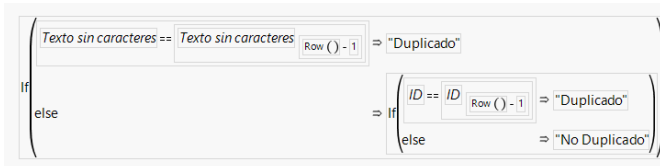


Fig. N°1. Regla de JMP para limpieza de datos

La polaridad se ensayó con la siguiente definición de un corpus básico de polaridad en la Figura N°2:

No	POSITIVO	NEGATIVO
1.	Mejor	Ratas
2.	Felicitaciones	Ladrón
3.	Felicidades	Mentira
4.	Compañeros	Malo
5.	Bien	Peor
6.	Bienvenido	Rechazo
7.	Confianza	Nunca
8.	Confiable	Vago
9.	Respaldo	Lame botas
10.	Bueno	Tirano
11.	Solidario	Corrupto
12.	Hermoso	Borrego
13.	Lindo	Chavista
14.	Éxito	Traición
15.	Ético	Traidor
16.	Aprobado	Inepto
17.	Adelante	Decepción
18.	Ganador	Perdedor

Fig. N°2. Corpus básico de polaridad

Se determina si es positivo con el valor de 1, neutro con el valor de 0 y negativo con el valor de -1.

Las siguientes son las métricas utilizadas para el estudio:

Número de seguidores, A esta métrica la hemos definido de tal manera que si nuestro personaje público tiene más de 200 seguidores en cada uno de sus tweets es aceptado por

parte de los twitteros en la red social.

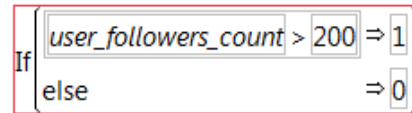


Fig. N°3. Regla de JMP de Número de seguidores

Aceptabilidad, esta métrica nos da a conocer que tan aceptado es el mensaje compartido por nuestro personaje por los twitteros en la red social.

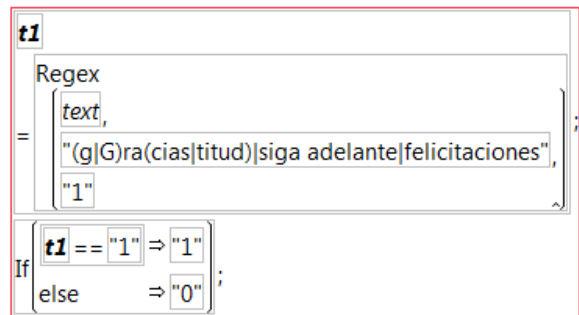


Fig. N°4. Regla de JMP de Aceptabilidad

Ratio de iteraciones /fans, esta fórmula nos permite determinar cuan relevante es una cuenta en relación al número de fans.

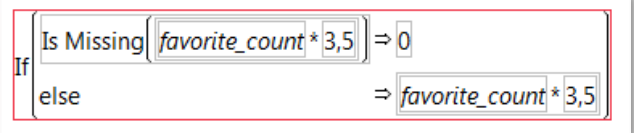


Fig. N°5. Regla de JMP de Valoración de favoritos

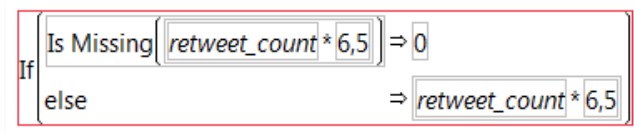


Fig. N°6. Regla de JMP de Valoración de retweets

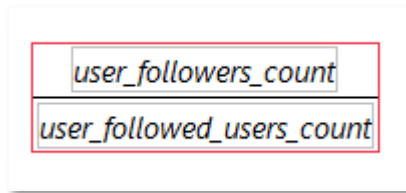


Fig. N°6. Regla de JMP de Ratio seguidores/seguidos

Valoración de Retweets + Valoración de favoritos

* Ratio Seguidores/Seguidos

Valoración de favoritos	Valoración de Retweets	Ratio Seguidores/Seguidos	Relevancia
0	0	0,952681388	0
0	0	0,952681388	0
0	6,5	11,757731959	76,425257732
84	149,5	124,77484183	29134,925568
0	3425,5	0,5596868885	1917,2074364
24,5	39	19,728401192	1252,7534757
0	0	111,77002967	0
28	208	17,157360406	4049,1370558
0	0	17,756722151	0
0	26	17,698209719	460,15345269
0	0	13,534482759	0
73,5	143	124,77484183	27013,753256
0	0	6,8653846154	0
0	962	0,023255814	22,372093023
0	6,5	2,1091674765	13,709588597
3,5	13	2,1091674765	34,801263362
0	253,5	2,1080467229	534,38984426
0	6,5	2,1080467229	13,702303699
0	0	2,1102106969	0
24,5	52	4,6952736318	359,18843284
0	97,5	4,6514143095	453,51289517
0	188,5	4,6514143095	876,79159734
38,5	84,5	4,6952736318	577,51865672
0	97,5	4,6514143095	453,51289517
7	13	4,6514143095	93,02828619
35	58,5	4,6952736318	439,00808458
0	110,5	4,6514143095	513,9812812
0	0	4,6514143095	0
10,5	39	2,1091674765	104,40379009
0	182	2,1080467229	383,66450357
0	260	2,1080467229	548,00214706

Fig. N°7. Regla de JMP de Relevancia y resultados

En lo referente a la polaridad del ejercicio se puede identificar que:

- 3% Negativo
- 92% Neutro
- 4% Positivo

Las opiniones sobre el personaje público son en su mayoría

neutras, definiendo que su accionar tampoco en negativo pero que tampoco da mucho que celebrar con opiniones positivas.

Al aplicar el score.sentiment, que es la función de predicción no supervisada se obtiene una puntuación positiva en los tweets, esto es comparando con un diccionario de palabras corpus.

Al establecer las métricas de relevancia a las polaridades positiva y neutral, se puede identificar lo siguiente:

- 14% Nada relevante
- 84% Poco relevante
- 0% Algo relevante
- 2% Normalmente relevante

Por medio del conteo de la métrica de relevancia, la información del cuadro indica que las opiniones sobre el personaje público fueron de poca relevancia, por los bajos niveles de retweets, y de poca marcación en favoritos.

V.CONCLUSIONES

1. Existen retos en el procesamiento del lenguaje natural que deben ser resueltos en la detección automática de la polaridad de un texto, esto es el corpus en español que entienda el léxico ecuatoriano o latinoamericano, este trabajo de investigación lo detecta e identifica una línea de estudio que puede ser trabajada en la comunidad de investigadores.

2. La definición de polaridad de sentimientos dentro de este estudio de opinión de un personaje público, determina indicios o tendencias al tipo de impacto, no tiene margen de error como una encuesta tradicional pero si provee un resultado que ayuda a reforzar la toma de decisiones.

3. La contribución del usuario en el lenguaje informal es determinante en una temporalidad cual puede variar en opinión, dichos textos deben ser preprocesados para obtener una data que aporte a un resultado relevante.

4. Los niveles de retweets y marcación en favoritos determinan la relevancia de una opinión, métrica utilizada para el estudio de una opinión en twitter, pueden ser también identificadas otras como ubicación geográfica, número de seguidores, etc.

REFERENCIAS

- [1] Borrás-Morell, Jose Enrique (2015) "Data Mining for Pulsing the Emotion on the Web" Colección: Methods in Molecular Biology Volumen: 1246 pp: 123-130
- [2] Chen, Xin; Vorvoreanu, Mihaela; Madhavan, Krishna(2014) "Mining Social Media Data for Understanding Students' Learning Experiences" IEEE TRANSACTIONS ON LEARNING TECHNOLOGIES Volumen: 7 Número: 3 pp: 246-259
- [3] Khan, Farhan Hassan; Bashir, Saba; Qamar, Usman (2014) "TOM: Twitter opinion mining framework using hybrid classification scheme" DECISION SUPPORT SYSTEMS Volumen: 57 pp: 245-257

- [4] Akaichi, Jalel; Dhouioui, Zeineb; Lopez-Huertas Perez, Maria Jose (2013) "Text Mining Facebook Status Updates for Sentiment Classification"
Conferencia: 17th International Conference System Theory, Control and Computing (ICSTCC) Ubicación: Sinaia, ROMANIA pp: 640-645
- [5] Colace, F.; De Santo, M.; Greco, L. (2014) "SAFE: A Sentiment Analysis Framework for E-Learning"
International Journal of Emerging Technologies in Learning Volumen: 9 Número: 6 pp: 37-41
- [6] Colace, Francesco; Casaburi, Luca; De Santo, Massimo (2015) "Sentiment detection in social networks and in collaborative learning environments" COMPUTERS IN HUMAN BEHAVIOR Volumen: 51 Número especial: SI pp: 1061-1067
- [7] Baldominos Gomez, Alejandro; Luis Minguez, Nerea; Cristina Garcia del Pozo, Ma (2015) "OpinAIS: An Artificial Immune System-based Framework for Opinion Mining" INTERNATIONAL JOURNAL OF INTERACTIVE MULTIMEDIA AND ARTIFICIAL INTELLIGENCE Volumen: 3 Número: 3 pp: 25-34

Metodología para procesos de inteligencia de negocios con mejoras en extracción y transformación de fuentes de datos

Santiago L.Morales C.¹, Mario Morales M¹, Ramon Rizo A.²
smorales@uce.edu.ec, mmoralesm@uce.edu.ec, ramon.rizo@gcloud.ua.es

Resumen— La inteligencia de negocios es un concepto que “integra como solución el almacenamiento y procesamiento de enormes cantidades de datos e información para transformarla en conocimiento y decisiones en tiempo real a través de una fácil explotación”. (Salinas La Rosa, 2010), así otro autor indica que “La Inteligencia de Negocios BI (Business Intelligence) es una herramienta bajo la cual diferentes tipos de organizaciones pueden soportar la toma de decisiones basadas en información precisa y oportuna; garantizando la generación del conocimiento necesario que permita escoger la alternativa que sea más conveniente para el éxito de la empresa”. (Rosado & Alveiro, 2010) En ese sentido el hecho de no tener claro un método para ser eficiente en la búsqueda de decisiones en inteligencia de negocios (BI) generan altos costos en el uso de recursos por lo cual la ciencia de la computación llegó a solventar esta problemática mediante una rama de esta que es la inteligencia artificial y en un sub campo como es el Aprendizaje Automático (A-A). En el presente trabajo se va a definir una metodología llamada M3S de IN, en la cual se determina un procedimiento adecuado a seguir para tener los mejores resultados en la extracción de información, para obtener mayor conocimiento y valor en la información; a su vez dicha metodología se sustentará en el algoritmo ID3 de árboles de decisión el cual se fundamenta de conceptos matemáticos como son la entropía y ganancia de información dentro del análisis de los mejores atributos que se pueden considerar para toma de decisiones.

Palabras clave: Inteligencia de Negocio, ID3, Aprendizaje automático, árboles de decisión, ETL.

ejecución de procesos de inteligencia de negocios (IN), que sirvan como estrategia para implantar este tipo de soluciones tecnológicas con éxito en empresas tanto del sector público como privado, mejorando principalmente los procesos de extracción y transformación de fuentes de datos heterogéneos con la ayuda de árboles de decisión.

OBJETIVOS ESPECÍFICOS:

- Desarrollar una metodología que permita establecer de manera óptima tanto en tiempo, como en recursos, soluciones de inteligencia de negocios (IN) con un alto grado de confiabilidad en el conocimiento obtenido.
- Contribuir con una nueva metodología formal en inteligencia de negocios (IN), para la generación de conocimiento en las organizaciones, particularmente en lo concerniente a manejo de indicadores relevantes de un determinado negocio o actividad de la vida.
- Usar aprendizaje automático con el objetivo ID3 de árboles de decisión dentro del proceso de extracción, transformación y carga de información. (ETC).

II. INTRODUCCIÓN

Los datos de origen para las aplicaciones de inteligencia de negocios (IN) provienen de varias plataformas heterogéneas, que son gestionadas por una variedad de sistemas operacionales y aplicaciones. Partiendo de la premisa de que los sistemas deben aprender y adaptarse a su entorno, el propósito de los procesos de extracción, transformación y carga de datos (ETC) en esta investigación es unir los datos de estas plataformas y transformarlos a un formato que contenga todas las clasificaciones de información necesarias, de tal manera que a partir de estos se tenga capacidad de decidir sobre todo el volumen de información actual y predecirla para llegar a un objetivo determinado (Rodallegas Ramos Erika, Torres González Areli, Gaona Couto Beatriz B, Gastelloú Hernández Erick, Lezama Morale Rafael As, V. O. S. 2010).

OBJETIVO GENERAL:

- Proponer, desarrollar y probar una metodología para la

III. DESARROLLO

Metodología

Dentro de la metodología en la fase de diseño y construcción ETC, se busca la opción de crear Árboles de Decisión (Quinlan. J.R, 1986), generados en base a experiencias concretas con los mejores atributos en lo que se refiere a ganancia de información que nos permitan testear nuevos datos y procesarlos para obtener resultados sobre hipótesis propuestas (Sempere, J. 2014).

El criterio principal es mejorar los resultados con la experiencia de información que contamos infiriendo suposiciones en hechos concretos (aprendizaje), lo cual nos permitirá realizar predicciones.

- Para esto utilizaremos el tipo de Aprendizaje Supervisado.
- Tenemos un conjunto de datos para los cuales buscamos

resultados.

- Queremos una o varias reglas que establezcan todas las opciones posibles.

IV. DISEÑO

Antes de proponer el diseño es bueno recordar que un árbol de decisión es un conjunto de condiciones organizadas en una estructura jerárquica, de tal manera que permite determinar la decisión final que se debe tomar siguiendo las condiciones que se cumplen desde la raíz del árbol hasta alguna de sus hojas Solarte Martínez, G. R. (2009). En base a lo descrito anteriormente planteamos sobre el Diseño ETC, los siguientes pasos para la aplicación de ID3 dentro de la metodología M3S

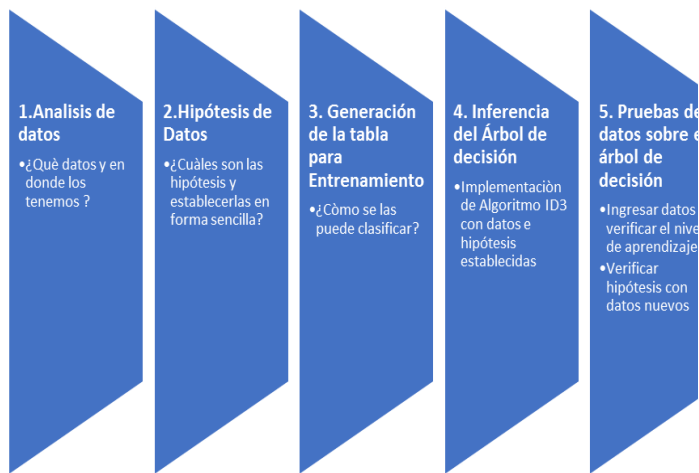


Figura 1: Fases de la metodología M3S, 2017

- 1. Análisis de Calidad de Datos:** Se procede con el análisis de los mejores datos que podamos tomar para un proceso de decisiones. Esto se obtendrá mediante uso de diccionarios de datos, estadísticas y conocimiento de expertos, para poder escoger los atributos más relevantes para el estudio.
- 2. Hipótesis de Datos:** El objetivo que tendrá esta parte metodológica es justamente encontrar una descripción del concepto objetivo mediante suposiciones generales a partir de hechos concretos de los datos en estudio y a raíz de estos clasificar la información para generalizar los casos nuevos.
- 3. Generación de la tabla para Entrenamiento:** Esta tabla es el producto final de los datos limpios y lo más refinados posible. Es producto también de la formalización de la o las reglas de negocio previamente establecidas con las cuales se han

logrado clasificar la información. A raíz de esta tabla podemos generar el proceso matemático y estadístico para la creación del árbol de decisión óptimo de acuerdo al algoritmo ID3 y a la objetividad que se haya dado en el análisis.

- 4. Inferencia del Árbol de decisión:** Para generar el árbol de decisión se utilizará inferencia inductiva que es aprendizaje supervisado. Este árbol es la aplicación del algoritmo ID3, el cual se basa en las reglas de negocio formalmente elegidas; del cual se prevé tener la mayor ganancia de información, y el mismo se encuentra detallado en la etapa de construcción. En forma simple para crear el indicado árbol se procederá de forma recursiva de arriba abajo. Utilizar en cada nodo el atributo “más importante” esto quiere decir que es aquel que mejor discrimina los ejemplos que han llegado hasta este nodo y por último clasificando los casos que siguen en el siguiente nivel.
- 5. Pruebas de datos sobre el árbol de decisión:** Una vez que contamos con el árbol de decisión, el cual con los cálculos de entropía nos genera la mayor ganancia de información; elaboramos un test para verificar que siempre se obtendrán los mejores resultados con nuevos datos que ingresen a evaluarse en el árbol. Se debe tomar en cuenta que, para algunos casos al probar nuevos datos, estos podrían hacer que cambie la clasificación y reglas de la tabla de entrenamiento, lo cual produciría una alteración en el árbol de decisión generado. Es por esto que al usar M3S, se debe tener claro que Inteligencia de Negocios lo podemos definir matemáticamente como:

$$IN = IN + APRENDIZAJE AUTOMÁTICO$$

Mitchell, T. (1997)

Formalizando esta metodología matemáticamente la planteamos a continuación:

- 1. Análisis de Calidad de Datos:**
Sea $f : X \rightarrow Y$, en donde:
 X son todos los casos posibles de entradas
 Y espacio de salidas
- 2. Hipótesis de Datos (H)**
Serán todas las $h \in H$

3. Generación de la tabla para Entrenamiento:

Si $f \in H$ entonces es Aprendizaje considerado

Si $f \notin H$ entonces no será considerado

4. Inferencia de Árbol de decisión:

Usar conocimiento a priori para asegurar que $H \subset f$ en donde la inferencia se lo realiza mediante:

1. Cálculo de la entropía

Dados:

- Un problema con dos clases (positiva y negativa)
- S, el conjunto de ejemplos.

$$Entropia(S) = \sum_i^k p_i \log_2 \left(\frac{1}{p_i} \right) \text{ con } k \text{ clases}$$

ECUACIÓN 1

2. Y también mediante la ganancia que se obtendrá.

$$Ganancia(S, A) = Entropia(S) - \sum_{v \in \text{valores de } A} \frac{|S_v|}{|S|} Entropia(S_v)$$

ECUACIÓN 2

5. Pruebas de datos sobre el árbol de decisión:

Utilizaremos medidas comunes de evaluación que son:

$$\text{promedio positivo verdadero} = \frac{TP}{TP + FN}$$

ECUACIÓN 3

$$\text{promedio positivo falso} = \frac{FP}{FP + TN}$$

ECUACIÓN 4

$$\text{precisión} = \frac{TP}{TP + FP}$$

ECUACIÓN 5

$$\text{corrección} = \frac{TP + TN}{TP + FP + TN + FN}$$

ECUACIÓN 6

Donde:

TP: Verdaderos positivos (clasificado: positivo, realidad: positivo)

TN: Verdaderos negativos (clasificado: negativo, realidad: negativo)

FP: falsos positivos (clasificado: positivo, realidad: negativo)

FN: falsos negativos (clasificado: negativo, realidad: positivo)

Es claro en esta propuesta que puede ser considerada para su difusión en el sector académico y a futuro por los estudiantes en el ámbito laboral en el cual se van a desarrollar, ya sea como emprendimientos personales o en relación de dependencia en cualquier institución pública o privada del país.

V. RESULTADOS EXPERIMENTALES

Se realiza actualmente una prueba de la metodología en la Empresa Eléctrica Quito, en donde se propuso experimentalmente ejecutar una campaña de uso de focos led, en base a diferentes criterios técnicos de expertos y la cual contó con una importante base de información de la facturación.

Se propusieron prototipos de tableros con el conocimiento generado, los mismos que se muestran en los siguientes gráficos:



Figura 2: Prototipos de Tableros.

De la propuesta ejecutada se generó un árbol de decisión que aplicando las reglas del negocio implementadas llegó a la siguiente estructura:

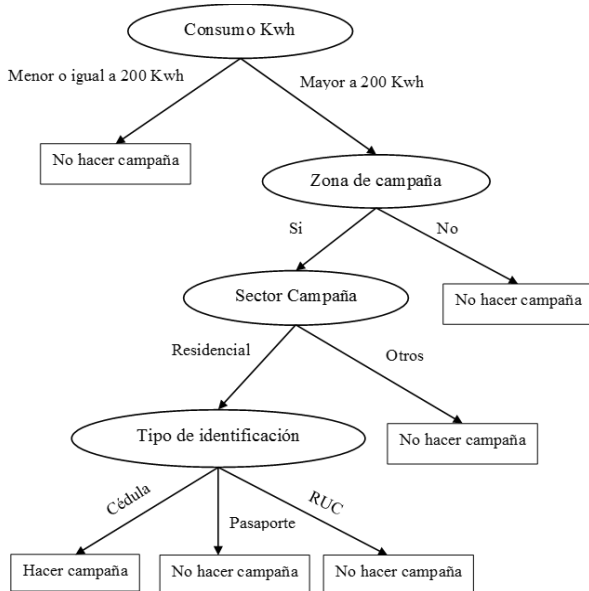


Figura 3: Arbol de Decisiones.Caso EEQ.

Finalmente se llega a concluir que de un total de 1.302.595 clientes analizados, a 87.506 de ellos que equivalen al 7% de la muestra deberán acogerse a la campaña de implementación de focos led en su domicilio.

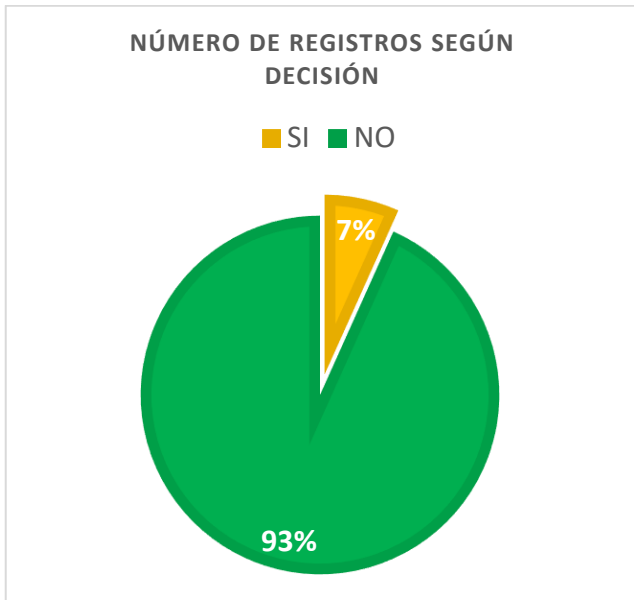


Figura 4: Hallazgo de Clientes

VI. CONCLUSIONES

- La presente metodología la consideramos una innovación que aportará en acortar los tiempos de proyectos en el área de Inteligencia de Negocios, logrando generar valor en las organizaciones.
- La eficiencia de los algoritmos de árboles de decisión depende en gran medida del conocimiento que nos entreguen los expertos de negocio, para generar una óptima tabla de entrenamiento (TE).
- La implementación de ID3 sobre TE, nos entrega un árbol de decisión que mantiene la mayor ganancia de información posible, que sin embargo está sujeto a cambios en base a nueva información que puede ingresar al sistema.
- Como resultados iniciales, se creó una Metodología dentro de ETC, en la cual se implementó ID3 para extraer información con mayor calidad de cualquier almacenamiento.
- Finalmente, basados en resultados experimentales iniciales, que deberán ser nuevamente validados, consideramos que los objetivos planteados pueden ser alcanzados en el tiempo previsto.

VII. REFERENCIAS

- Salinas La Rosa, A. (2010). *Inteligencia de Negocio. Auditoría y Control. Prototipo de Herramienta de Calidad* (Vol. 1). (D. D. INFORMATICA, Ed.) Leganes, España: UNIVERSIDAD CARLOS III DE MADRID. Obtenido de <https://core.ac.uk/download/pdf/29402111.pdf>
- Sempere, J. (2014). *Aprendizaje de árboles de decisión*. Valencia. Universidad Politécnica de Valencia.
- Mitchell, T. (1997). *Machine Learning*. McGraw Hill.
- Rodallegas Ramos Erika, Torres González Areli, Gaona Couto Beatriz B, Gastelloú Hernández Erick, Lezama Morale Rafael As, V. O. S. (2010). *Minería de datos: predicción de la deserción escolar mediante el algoritmo de árboles de decisión y el algoritmo de los k vecinos más cercanos. Recursos Digitales*. Retrieved from http://ccita2010.utmetropolitana.edu.mx/recursos/Recursos_digitales.pdf#page=34
- Quinlan. J.R (1986), *Induction of Decision Trees*. Recuperado de http://www.scielo.br/scielo.php?pid=s1807-17752013000200389&script=sci_arttext&tlng=pt#b12
- Solarte Martínez, G. R. (2009). *Técnicas de clasificación y análisis de representación del conocimiento para problemas de diagnóstico*.

Análisis de la respuesta espectral en plantaciones de banano, a través de la relación de los micro elementos del suelo y los índices de vegetación, en la finca blanca américa, provincia del Guayas, cantón San Jacinto de Yaguachi.

Lizbeth Tipantuña¹

liz16020691@gmail.com

¹Pontificia Universidad Católica del Ecuador

Resumen. - La presente investigación se la realizó en la Finca Blanca América, ubicada en la Provincia del Guayas, cantón San Jacinto de Yaguachi, la cual tiene una superficie de 129.26 ha, dedicada a la producción y comercialización de Banano, investigación que es la primera que con tecnologías geomáticas y de sensores remotos se realiza en el país.

El propósito de la investigación se enfoca en analizar la respuesta espectral del cultivo de Banano (*Musa spp.*) mediante la utilización de diferentes índices de vegetación, aplicados a una imagen de satélite Pleiades de 0,5m de resolución del área de estudio, con la finalidad de la detección de plagas y enfermedades.

Con la imagen del satélite Pleiades, se procedió a su identificación y corte de acuerdo al límite de la propiedad, para posteriormente evaluar tres índices de vegetación: Índice de Vegetación de Diferencia Normalizada NDVI, Índice Mejorado de Vegetación EVI, y el Índice Transformado de Vegetación SAVI, los cuales fueron utilizados como estimadores de diferentes cambios en la cobertura del cultivo de Banano.

Así mismo para correlacionar información edáfica y la satelital, se tomó siete muestras, tanto de suelo como foliares, las mismas que fueron enviadas al laboratorio para su respectivo análisis.

Se aplicaron técnicas de análisis espacial como el Método de interpolación de Kriging, en base a los resultados reportados por el laboratorio para así obtener mapas de fertilidad de los macro elementos Nitrógeno, Fósforo y Potasio, los cuales fueron correlacionados con los diferentes índices obtenidos anteriormente.

El análisis de las características de los índices de vegetación y el análisis de los elementos químicos permitió identificar valores y/o rangos que muestren el estado del cultivo y a su vez dar alertas sobre posibles sitios en donde pueda existir algún tipo de ataque de alguna plaga u enfermedad o algún tipo de deficiencia de algún elemento químico necesario para la planta.

En este sentido la aplicación de la metodología busco dar un seguimiento a la problemática de la sanidad del cultivo dentro de la propiedad, relacionando el estado fenológico del cultivo con la información de la imagen satelital. obtenida, permitiendo así tomar las mejores decisiones en cuanto a un mejor manejo del cultivo optimizando recursos e insumos dentro del manejo del cultivo de banano.

Palabras Claves: propuesta metodológica, tecnologías geomáticas, sensores remotos y SIG.

I. Introducción.

Las características fenológicas que tienen los diferentes cultivos permiten identificar el estado de la cobertura vegetal, siendo una de ellas el verdor que posee la vegetación ya que este es un indicativo de una buena actividad fotosintética y por ende un indicativo de una buena salud de la cobertura vegetal de la planta.

A lo largo de las últimas décadas varios investigadores han propuesto diversos índices espectrales, que según la característica que se busca evaluar, ha generado un gran número de expresiones aplicadas a distintos estudios en la vegetación, los mismos que se emplean para cuantificar el comportamiento fenológico de las coberturas vegetales. Dichos índices se basan en la respuesta espectral de la vegetación, con datos que se obtienen a partir de sensores remotos, los cuales se encargan de recoger información sobre la radiación que refleja la superficie terrestre y posteriormente la almacena en diferentes bandas del espectro electromagnético, dependiendo de su longitud de onda.

El conjunto de estas bandas proporciona imágenes o matrices de datos en los que cada

pixel contiene información de reflectancia para una zona determinada de la superficie. Para la presente investigación se consideró la constelación de los satélites Pléiades 1A y 1B, éstos satélites son totalmente idénticos suministran productos ópticos de alta resolución con una capacidad de revisita diaria sobre cualquier punto del globo, además es idónea para obtener información en tiempo récord con una capacidad de adquisición inferior a 24 horas sobre cualquier punto del planeta que permitiría obtener información en caso de crisis o de catástrofe natural. La utilización de los diferentes índices de vegetación y en especial de Índice de Vegetación de Diferencia Normalizada NDVI, ha sido ampliamente utilizado en la identificación del uso del suelo, en la caracterización espectral de cultivos, en la determinación del vigor de los mismos, en aportes a la investigación de plagas y/o enfermedades. Existen otros índices de vegetación que también fueron utilizados en la presente investigación, y que aportan con la información y datos relativos a la biomasa/verdor de la vegetación y/o humedad y que se han aplicado a la evaluación de cultivos son: el Índice de Vegetación de Ajustado al suelo (SAVI) por sus siglas en inglés que tiende a minimizar el brillo del suelo, introduciendo una variable L

[1]; y, el Índice de Vegetación Mejorado (EVI), aplicado por su información en la banda azul del espectro, y que puede ayudar a corregir señales de fondo e influencias atmosféricas [2].

En este contexto, el objetivo de la investigación se enfocó en analizar los diferentes índices de vegetación, además de una correlación con los

análisis de suelo y foliares como una herramienta para identificar plagas, enfermedades además de deficiencias de algún elemento químico en el cultivo de banano.

II. DESARROLLO

Metodología.

La metodología consistió en la recopilación, sistematización y análisis de información básica. Se recopiló datos básicos e informativos de la zona de Estudio (temperatura, precipitación, tipo de suelo, y vegetación), además de información de varias fuentes que contengan el uso de los Índices de vegetación (NDVI – EVI - SAVI), los datos históricos de las principales plagas y enfermedades una zona piloto del cantón San Jacinto de Yaguachi sector Tres Postes, Provincia del Guayas y finalmente la recopilación y análisis de imágenes satelitales Pléiades obtenidas de AIRBUS.

Levantamiento de información de campo: Se identificó siete puntos de muestreo, se realizó el levantamiento de una ficha de campo en la finca se procedió con la obtención de coordenadas dentro de cada parcela con presencia de plagas y enfermedades o cualquier otro factor.

El procesamiento de la información del estudio: se realizó un análisis de las respuestas espectrales reflejadas de la planta del Banano (NDVI), además del procesamiento de modelos matemáticos generados para el análisis espectral. Y finalmente se realizó la correlación entre la respuesta de reflectancia del cultivo en la imagen satelital y los datos tomados en el campo mediante el álgebra de mapas. Ver Figura N°1

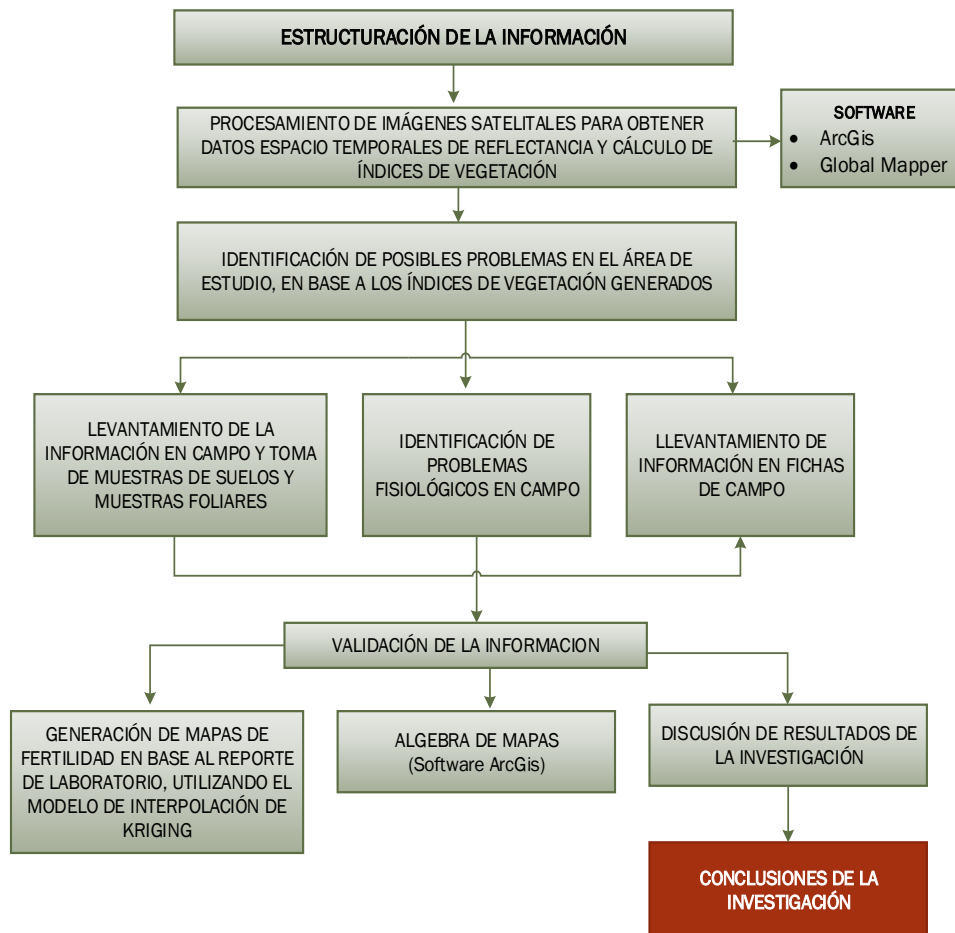


Figura N°1 Diagrama de flujo de la metodología a aplicar para la utilización de diferentes algoritmos de teledetección

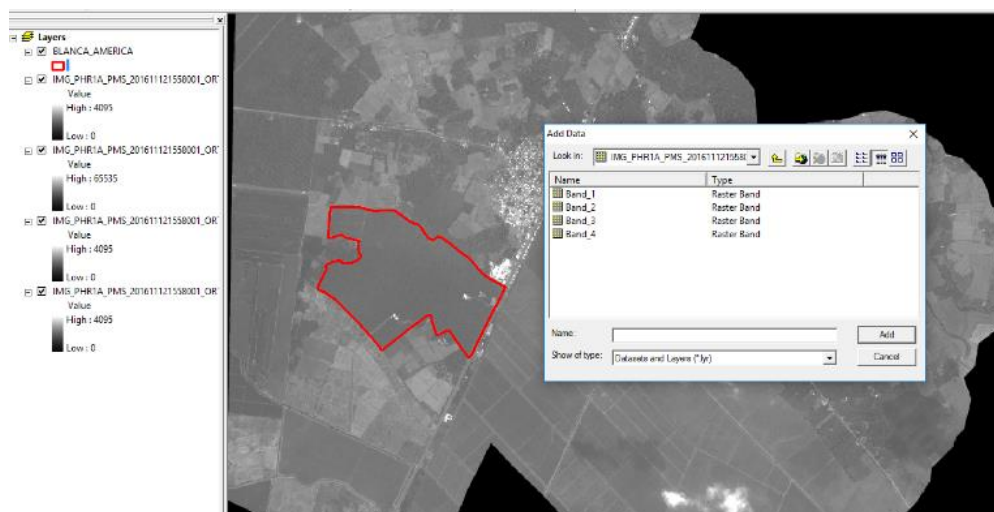


Figura 2. Visualización de Bandas de la Imagen Pleiades

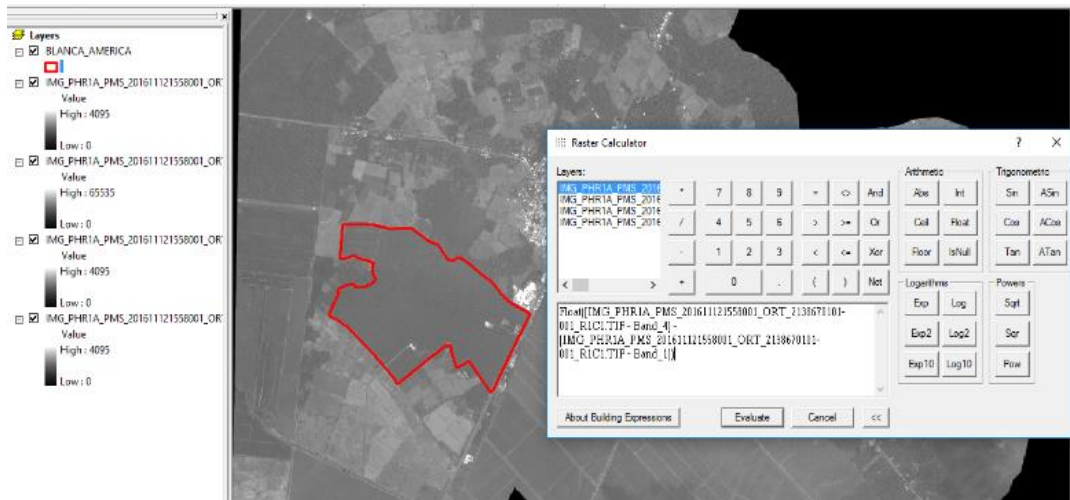


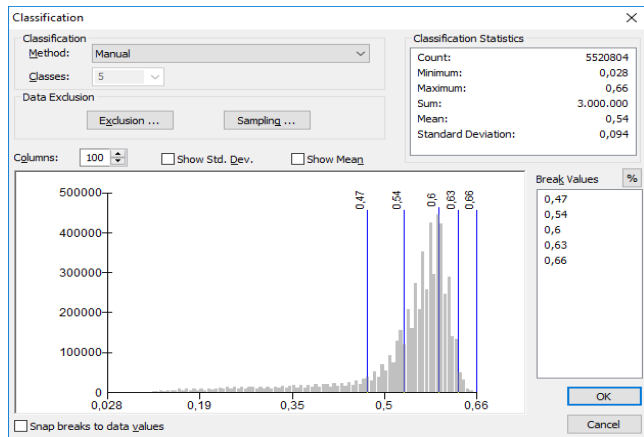
Figura 3. Aplicación de la herramienta “Raster Calculator”

III. RESULTADOS.

Para proceder con el cálculo de los diferentes índices de vegetación se despliega la imagen con las cuatro bandas que la conforman: rojo, verde, azul, infrarrojo cercano de tal manera que con la ayuda de la “Calculadora Raster” ingresar los diferentes arreglos matemáticos para el cálculo de los mismos. Cabe mencionar que dicha herramienta permite crear y ejecutar una expresión de Algebra de Mapas y que a su vez genera como salida un Raster, en este caso de los tres índices analizados (NDVI, SAVI y EVI). La herramienta Calculadora ráster permite crear y ejecutar una expresión de Álgebra de mapas que generará como salida un ráster.

Índice de diferencia normalizada de vegetación NDVI.

El cálculo del NDVI, se generó de acuerdo con el siguiente modelo matemático:



$$NDVI = \frac{Banda\ 4\ (infrarrojo\ cercano) - Banda\ 1\ (Roja)}{Banda\ 4\ (infrarrojo\ cercano) + Banda\ 1\ (Roja)}$$

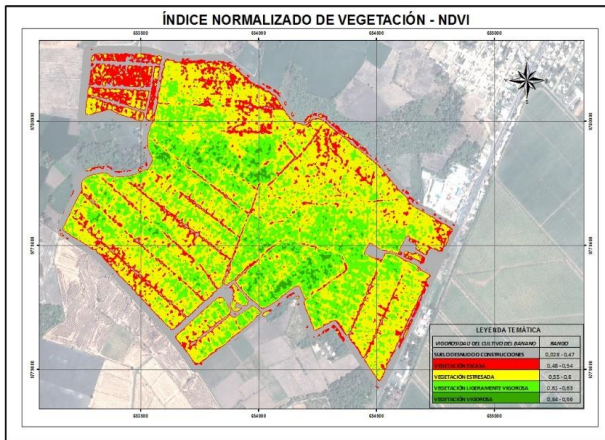


Figura 4. Índice de diferencia normalizada de vegetación (NDVI).

Figura 5. Clasificación en forma de histograma. (NDVI)

Los valores calculados del NDVI, se encuentran dentro del rango de 0.028 a 0.66, rango que posteriormente fue reclasificado en cinco niveles diferenciando los diferentes estados de la vegetación, evidenciándose así que las zonas con una mejor actividad fotosintética y por ende mayor vigor, se encuentran en el rango de 0.61 a 0.66, en tanto que los valores entre 0,02 a 0,54 hacen referencia a categorías de vegetación escasa y suelo desnudo o construcciones

Índice mejorado de vegetación EVI

El cálculo del EVI se generó de acuerdo con el siguiente modelo matemático:

$$Evi = \frac{2,5 * (banda\ 4 - banda\ 1)}{[1 + banda\ 4 - banda\ 1 - 7,5 / banda\ 3]}$$

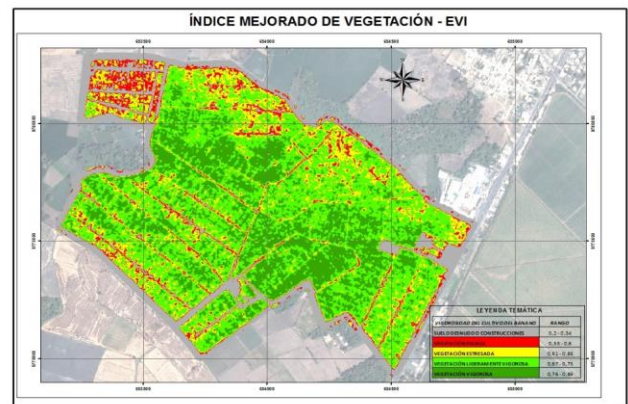


Figura 6. Índice Mejorado de vegetación EVI y rangos de clasificación.

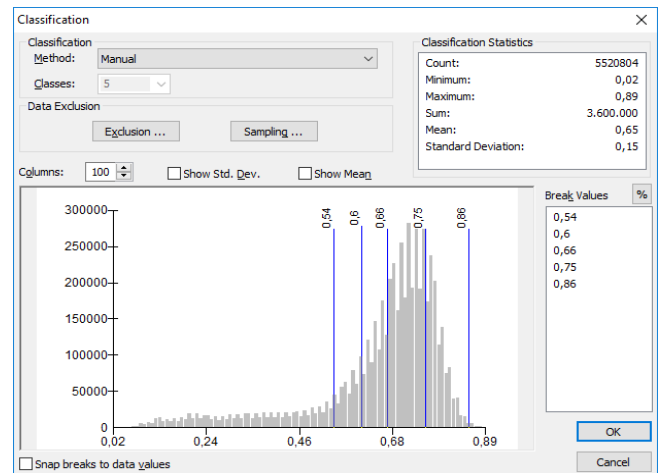


Figura 7. Clasificación en forma de histograma. EVI.

Al aplicar el índice EVI se evidencia una mejora sobre el NDVI, este índice nos arroja valores entre 0,02 a 0,86. En la imagen podemos decir que los valores acumulados oscilan entre 0,75 y 0,86; esto nos permite tener una mejor

discriminación en cuanto a la vegetación. En este entorno los valores que se encuentran en el rango de los 0,75 y 0,86 hacen referencia a valores con una fuerte actividad fotosintética, demostrando así una vegetación sana. Los valores bajos 0,02 a 0,6 hacen referencia a las categorías vegetación escasa y suelo desnudo o construcciones.

Índice transformado de vegetación SAVI

El cálculo del SAVI se generó de acuerdo con el siguiente modelo matemático:

$$\text{SAVI} = [(1 + 0,5) * (\text{Banda 4} - \text{banda 1})] / [(\text{Banda 4} + \text{banda 1} + 0,5)].$$

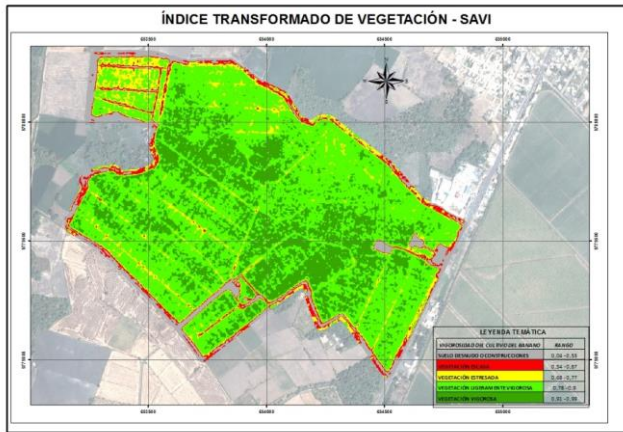


Figura 8. Índice Transformado de Vegetación SAVI

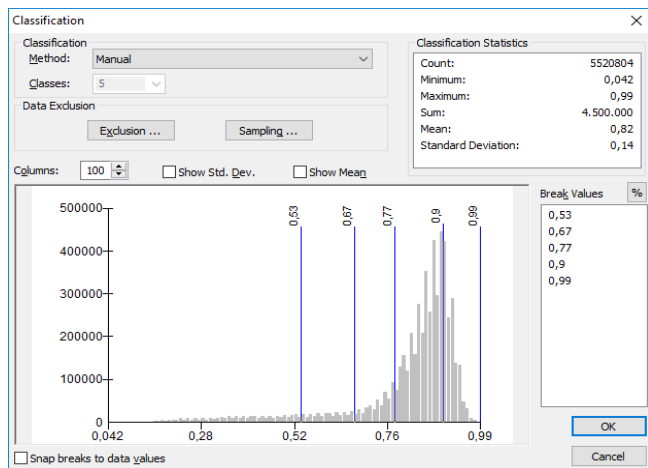


Figura 9. Clasificación en forma de histograma. SAVI

El índice SAVI muestra una distribución normal, sus valores de ND (niveles digitales o valores del pixel), para este caso observamos valores en la imagen que se encuentran entre 0,04 a 0,99, posteriormente fue reclasificado en cinco niveles diferenciando los diferentes estados de la vegetación, evidenciándose así que las zonas con una mejor actividad fotosintética y por ende mejor estado de salud se encuentran

en el rango de 0.78 a 0.99, en tanto que los valores entre 0,04 a 0,67 hacen referencia a categorías de vegetación escasa y suelo desnudo o construcciones

Análisis espacial para muestras de suelo y muestras foliares

Como se mencionó anteriormente se utilizó técnicas de análisis espacial como el Método de Interpolación Kriging, el cuál es un procedimiento geoestadístico avanzado que genera una superficie estimada a partir de un conjunto de puntos dispersos con valores z [3].

Este proceso Geoestadístico se puede resumir en tres etapas:

- Análisis exploratorio de los datos
- Análisis estructural
- Predicción de las variables

Para de esta manera cumplido este proceso obtener los diferentes mapas de fertilidad.

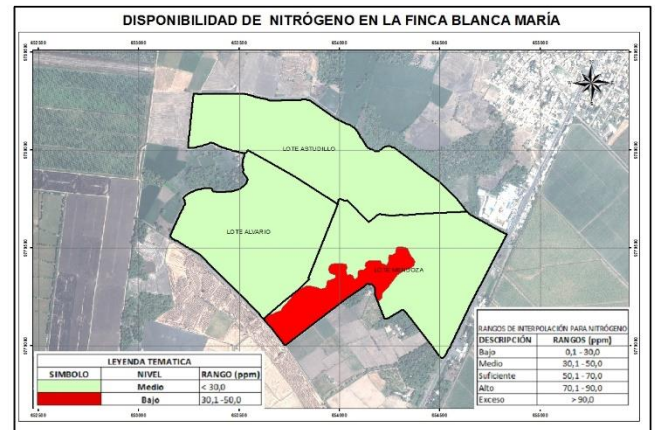


Figura 10. Disponibilidad de N en el suelo.

El contenido de nitrógeno en los suelos varia en un amplio espectro. El nitrógeno tiende a incrementarse al disminuir la temperatura de los suelos y al aumentar las precipitaciones atmosféricas. Si el suelo tiene deficiencia de nitrógeno, las plantas se vuelven altas, débiles y pálidas [4].

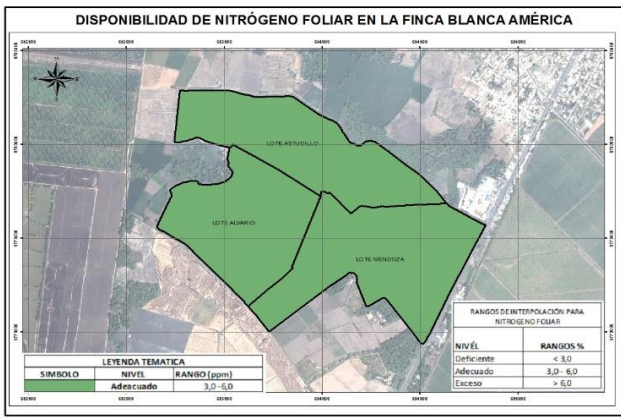


Figura 11. Disponibilidad de N foliar.

Interpretado los resultados de laboratorio para muestras foliares y generado el mapa para el contenido de Nitrógeno, se determinó que el contenido de este elemento en las hojas es adecuado lo cual se vio evidenciado en el follaje del cultivo.

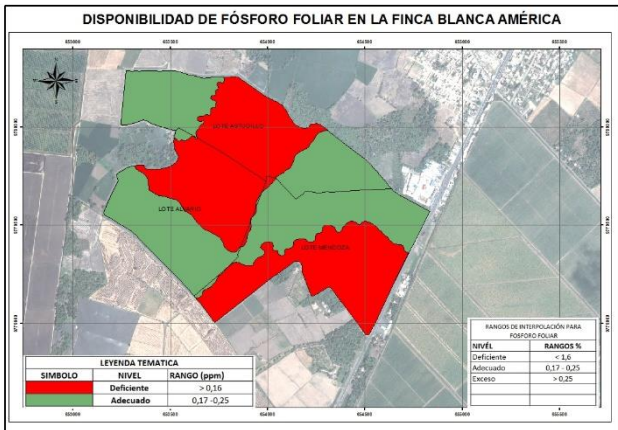


Figura 12. Disponibilidad de P en el suelo

El elemento fósforo se encuentra en niveles altos, seguramente por una fertilización constante, la cual podría suspenderse y con nuevos análisis en futuro evaluar si se requieren nuevas aplicaciones de fertilizante para evitar bajones en la disponibilidad de este importante elemento.

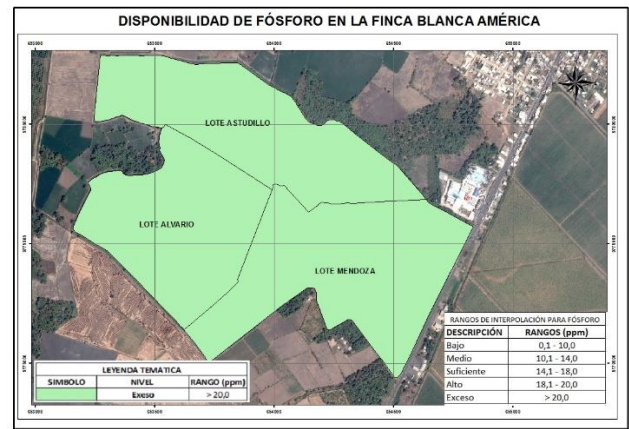


Figura 13. Disponibilidad de P foliar.

Interpretado los resultados de laboratorio para muestras foliares y generado el mapa para el contenido de Fosforo, se determinó que existen zonas en el que el contenido de este elemento es deficiente, para lo cual es necesario incorporar este elemento al cultivo con el propósito que la deficiencia de este elemento pueda producir un crecimiento inadecuado de la parte foliar e incluso un decrecimiento de la masa radicular de la planta.

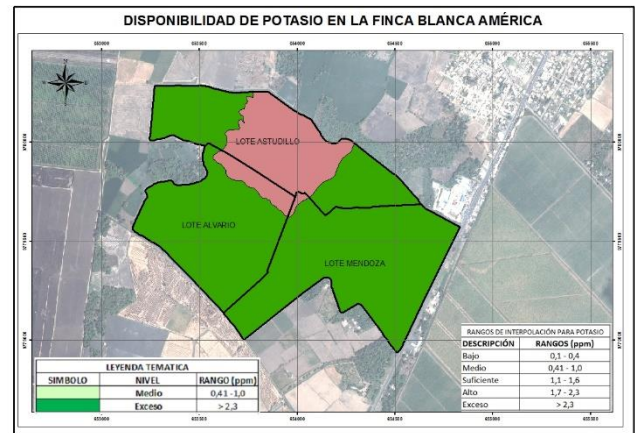


Figura 14. Disponibilidad de K en el suelo.

El potasio (k) es uno de los macronutrientes esenciales más importantes que permiten el funcionamiento de sistemas agropecuarios. Cumple funciones vitales en la fisiología vegetal y por lo tanto su deficiencia origina importantes desgastes en el rendimiento y/o calidad de los cultivos [4]. El Potasio se encuentran en altas cantidades, pero entre ellos desbalanceados, siempre se requiere que estén en una relación aproximada de Ca-Mg-K de 70-20-10, pero en este caso el elemento Potasio se encuentra bajo de esta relación, es de decir, que a pesar de existir altas cantidades son dominadas por los cationes Calcio, lo que crea una aparente

deficiencia de K, por lo que es necesario aplicar K para suplir esa deficiencia.

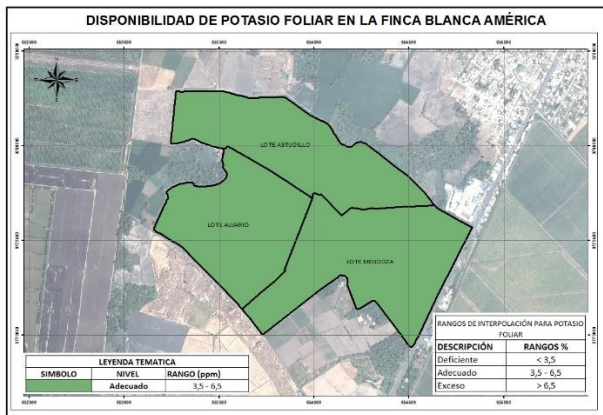


Figura 15. Disponibilidad de K foliar.

Interpretado los resultados de laboratorio para muestras foliares y generado el mapa para el contenido de potasio, se determinó que el contenido de este elemento en las hojas es adecuado lo cual se vio evidenciado en el follaje del cultivo. Además, con la información reportada por el laboratorio se generó mapas de Fertilidad los cuales sirvieron para contrarrestar con la información obtenida de la generación de los diferentes índices de vegetación.

IV. CONCLUSIONES

1. Los índices de vegetación generados para el cultivo de Banano arrojaron resultados satisfactorios, en especial el Índice de Vegetación Normalizado (NDVI) y el Índice Mejorado de Vegetación (EVI), evidenciándose que la mejor respuesta espectral la obtuvo el Índice Mejorado de Vegetación (EVI) ya que evidencio una mejor discriminación y clasificación.
2. La validación de la información reportada para muestras foliares y muestras de suelo y su correlación con la respuesta espectral de la planta fueron satisfactorias en algunos casos por ejemplo el fósforo foliar con el EVI, el fósforo Suelo con SAVI y el potasio Suelo con SAVI tiene una correlación directa.
3. La simple percepción o medición de los datos espectrales no es suficiente para constituirse en una metodología de análisis, para esto es necesario llevar a cabo un proceso de calibración de los datos, que consiste en relacionar los datos espectrales con la o las variables que se están estudiando.
4. La metodología propuesta fue apropiada para la evaluación de procesos en la identificación de

plagas y enfermedades, ya que se dieron resultados positivos y plantados en los objetivos antes mencionados.

REFERENCIAS

- [1] Huete A.R., Liu, H.Q., Batchily, K., van Leeuwen, W. (1988): A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sensing of Environment.
- [2] Holzman, M., Rivas, R., Piccolo, M. (2013). Estimating soil moisture and the relationship with crop yield using surface temperature and vegetation index. Intern Journal of Applied Earth Observation and Geofomation, 28:181–192.
- [3] Henao R. 2005. Introducción a la geoestadística, Teoría y Aplicación. Universidad Nacional de Colombia. 5 – 42 p.
- [4] Bertsch, F. (2003): Absorción de nutrimentos por los cultivos. San José, Costa Rica, Asociación Costarricense de la Ciencia del Suelo. 307 p.

Sistemas de control distribuido para el proceso de generación de vapor en una planta termoeléctrica

Almeida Clavijo, Eduardo¹; Yáñez-Ortiz, Verónica²
 walter.almeida@pucese.edu.ec, veronica.yanez@pucese.edu.ec
 Pontificia Universidad Católica Sede Esmeraldas – Ecuador

Resumen— La evolución de los procesos en plantas industriales va ligada con el desarrollo tecnológico, su manejo ha cambiado paulatinamente de un modo manual a un modo automático, este progreso lento está justificado porque se ha debido priorizar la seguridad operativa de los nuevos sistemas, con el fin de garantizar la creación de procesos productivos confiables y eficientes. En este proyecto se planteó evidenciar cómo la implementación de un sistema de control distribuido (Delta V), apoyado con tecnologías avanzadas, puede mejorar la operación de una caldera para producir vapor en una planta de generación termoeléctrica. El cambio del sistema de control se lo hizo recolectando los datos generados por los equipos de campo, cuya señal (analógica o digital) fue enviada a los módulos de acoplamiento, encargados de cambiar la información a lenguaje de programación del controlador lógico programable (PLC), que permite ser configurado en línea, el cual estableció la comunicación con la interfaz hombre-máquina (HMI), permitiendo que el operador visualice el proceso de producción de vapor en tiempo real desde su computador. La implementación de este sistema permitió mejorar la confiabilidad del manejo de la planta ya que al acceder a más información del proceso se puede tomar decisiones efectivas y a tiempo; la generación de un historial del comportamiento de las variables de los diferentes equipos también contribuyó a realizar cambios oportunos a través de mantenimientos predictivos.

Palabras clave: sistemas de control distribuido, generación de vapor, planta termoeléctrica, HMI, PLC, módulos de acoplamiento, Delta V.

I. INTRODUCCIÓN

La creación de la máquina de vapor y la producción en línea, provocan la aparición de sistemas automáticos, que tienen como fin facilitar el control de procesos industriales; el avance de la tecnología ha fortalecido estos sistemas, haciendo que el hombre pueda manejar los procesos productivos desde su computador, tomando decisiones y realizando cambios apoyado en análisis de datos de los equipos de campo.

Los sistemas antiguos trabajaban con modelos analógicos que tenían ciertas limitaciones, en cuanto a la cantidad de información que manejaban, el volumen que ocupaban y los recursos para su funcionamiento, espacio, cables, temperaturas de enfriamiento [1]. La figura 1 muestra un controlador con módulos conectados para un sistema de 4-20 mA empleado en la caldera antes de implementar el sistema Delta V.

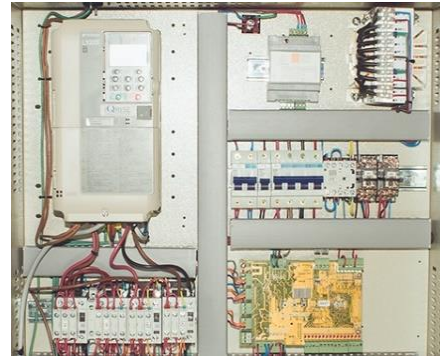


Figura 1. Sistema de control analógica de la caldera

Fue cuestión de tiempo para que los sistemas cambiaran de rumbo hacia las nuevas tecnologías, tal es el caso de las empresas Generadoras de Energía Termoeléctrica en el Ecuador, para citar un ejemplo, estas empresas que fueron construidas a inicios de los años 80s, tenían en su tiempo la última tecnología, luego se vieron en la necesidad de migrar a otra tecnología como la de los PLCs, la que ha evolucionado dando lugar a los sistemas de control distribuido o también llamados sistemas SCADA (Supervisión, Control y Adquisición de Datos), uno de los cuales es el Sistema Delta V.

Por otro lado, han mejorado potencialmente las comunicaciones entre los PLCs y el HMI, en la Figura 2 se aprecia como se ha dado esta evolución mediante redes que van desde las tecnologías RS232, RS485, modbus y Ethernet / IP; dando más confiabilidad al sistema [2].

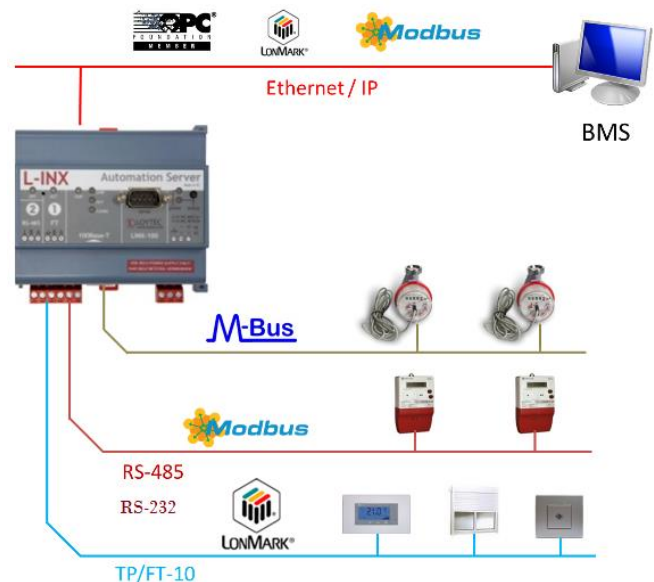


Figura 2. Comunicación entre PLCs y HMI

Fuente: <http://e-controls.es/es/notas-de-aplicaci%C3%B3n/pasarela-multiprotocolo-lonworks-modbus-m-bus-y-opc-para-automatizaci%C3%B3n-de>

En el proceso de generación de energía termoeléctrica, el caldero cumple la función de producir vapor para mover la turbina que origina la electricidad. En este proyecto, se incorpora el sistema de control distribuido Delta V a la caldera mediante el cambio de la instrumentación básica de presión, temperatura, flujo, concentración de oxígeno y conductividad del agua del condensador. Este cambio incluye la renovación completa de tarjetas electrónicas con PLCs y módulos de acoplamiento y la implementación del HMI que, a través de la configuración adecuada de datos, presenta la información relacionada con el proceso de generación de vapor (variables y equipos) en pantallas, facilitando la operación del caldero.

Los transmisores, encargados de cambiar las mediciones de las variables físicas (presión, temperatura, concentración, conductividad) a una señal eléctrica receptada por el módulo de acoplamiento [3]; tienen diferentes modos de comunicarse al Delta V, entre los principales se tiene:

- 4-20 mA, son los más básicos, presentan sólo los datos del equipo, no se incluye en un lazo de control, muestran únicamente la información del transmisor. Se comunican con el Delta V, por medio de Módulos de 4-20 mA de entrada.
- HART, transductor remoto direccionable de alta velocidad, puede comunicar una sola variable usando una señal analógica de 4-20 mA, mientras comunica también información agregada sobre una señal digital como la de temperatura del equipo [4], este transductor no está incluido en un lazo de control (transmisor-control-válvula). Se comunican con el Delta V, por medio de Tarjetas 4-20 mA-HART.
- MODBUS, es un bus de datos, en el cual la información llega en bloque, por ejemplo, las temperaturas de tubos del caldero, que son recolectadas por un PLC diseñado para ello, luego estos datos son enviados por medio de comunicación MODBUS al Módulo H1 específico para este tipo de comunicación del sistema Delta V [5].
- FIELDBUS, su función es coleccionar y entregar grandes cantidades de información, no sólo variables del proceso y señales de control, sino también, otros tipos de datos de los instrumentos y del proceso [6]. También transporta la información al módulo H1, que incluye: la señal 4-20mA, temperatura, densidad, estado del instrumento, alarmas.

Los datos recogidos en los módulos pasan al controlador, que es el cerebro del proceso, en donde se ubican las estrategias del Delta V, que constituyen básicamente el programa de control de la caldera. En las estrategias se consideran: funciones analógicas de entrada y salida, función PID (proporcional integral derivativo), bloques de alarmas, diseños lógico matemáticos [7]. Esta información pasa vía Ethernet al computador que maneja el operador de la caldera

mediante la interfaz dispuesta en pantallas que permiten visualizar los datos del proceso que le facilitarán, previo un análisis, a tomar decisiones que ayuden a mantener la eficiencia en la producción de vapor.

En la actualidad, muchos de los procesos en las industrias se han modernizado, empleando sistemas de control distribuido desarrollado por diversas casas comerciales, tales como Honeywell, ABB o Emerson (Delta V), que transforman la relación operador - proceso productivo, convirtiéndola en una forma de optimizar recursos mediante: el manejo predictivo de las variables de operación [8]; el control del proceso en tiempo real evitando su interrupción; la escalabilidad del proceso incorporando, en línea, nuevas fases de control; la facilidad de manejo del sistema mediante instrucciones de orden y acción presentadas en pantallas que el operador puede manipular fácilmente.

En este artículo se explica la metodología utilizada, el diseño del sistema de control, los resultados obtenidos y finalmente se presentan las conclusiones.

II. DESARROLLO

1. Metodología

La metodología empleada consistió en determinar las variables presentes en la producción de vapor de la caldera, con esta información se escogieron los instrumentos necesarios para medir estas variables. Los datos generados en estos instrumentos fueron transmitidos a los módulos de acoplamiento a través de cables ubicados en tuberías previamente instaladas. En los módulos de acoplamiento se transformó la señal analógica-digital de los diferentes instrumentos, la información generada fue compartida con los controladores MD plus, mediante estrategias propias del Delta V, los datos se presentaron en el computador del operador a través de pantallas diseñadas para facilitar el control del proceso. En la Figura 3 se puede observar los módulos y controladores MD plus que permiten la automatización del caldero.



Figura 3: Módulos y controladores del Delta V
Fuente: Emerson-Delta V

III. DISEÑO

El sistema Delta V, se lo considera un sistema distribuido, ya que tiene varias partes desde los programas de mando y configuración del HMI; la red de comunicación Ethernet, por medio de la cual se comunican entre computadoras; los controladores y módulos, los cuales procesan la información que proviene de los sensores o transmisores de campo.

En lo referente al software, el Delta V tiene programas básicos, los cuales son: el *Delta V configurator*, en el cual se configuran los usuarios, la distribución de los controladores y computadoras; el *Delta V-AMS*, que contiene la base de datos de los transmisores y equipos de campo, sus configuraciones y datos; el *Delta V- Explorer*, a través del cual se puede navegar por los distintos programas que posee el Delta V e identificar cada elemento de campo; el *Delta V-Histórico*, que permite ver las gráficas de los datos en línea y tomar medidas en caso de perturbaciones de equipos, la pantalla que muestra estas gráficas se aprecian en la Figura 4; el *Delta V-Estrategias*, con el que se programa el cerebro del proceso, en este se interconectan los módulos analógicos, digitales, PID para el control de válvulas de control [9].

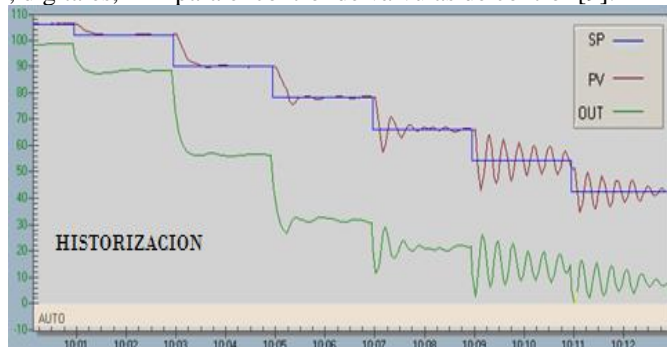


Figura 4. Histórico de una variable
Fuente: Emerson-Delta V

Un lazo de control se conforma por el elemento sensor, luego el controlador y finalmente el elemento actuador. Para el caso del Delta V, el transmisor es el elemento sensor el cual puede ser de 4-20mA, HART y Fieldbus, la señal generada llega a su respectivo módulo y se enlaza con el Delta V a través de su estrategia, en la cual hay un bloque PID que controla el elemento final, que es la válvula que permite el paso de vapor.

Algunos transmisores no controlan el proceso, solo sirven para ver los valores de las variables de campo (presión, temperatura, flujo), y de esta manera tomar decisiones en el caso de presentarse algún evento que modifique su comportamiento habitual. Las señales vienen de los transmisores 4-20 mA o HART, o en bloque de datos por medio de comunicación MODBUS. Cada elemento tiene su respectivo módulo, por ejemplo, un transmisor 4-20 mA requiere un módulo 4-20 mA de entrada; un transmisor HART, requiere un módulo HART; y un bus de datos MODBUS, requiere un módulo H1, específico del sistema Delta V.

También existen elementos WIRELESS, que se comunican por señal de radio a una frecuencia que no afecta a los sistemas localizados de radio enlace. Además, existe una

antena para recibir la señal y enviarla a su módulo WIRELESS respectivo.

En la Figura 5 se puede apreciar los tipos de comunicación entre los instrumentos de campo y el HMI.

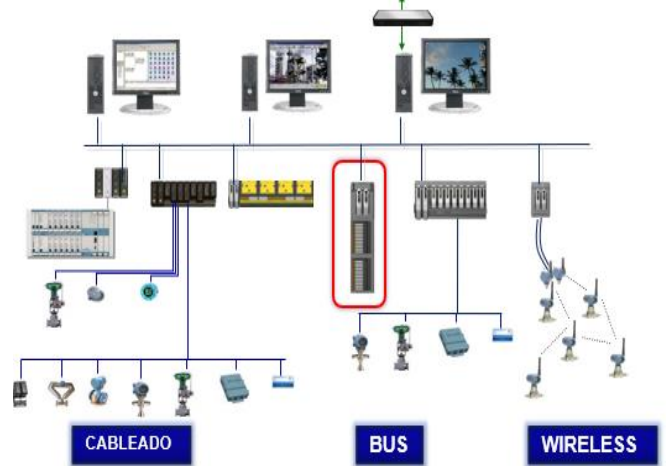


Figura 5. Comunicación entre equipos de campo y el HMI
Fuente: Emerson-Delta V

Cada dato que entra a la estrategia del DELTA V, muestra su valor en la pantalla HMI. Como se muestra en la Figura 6, en las pantallas del Delta V se pueden colocar los controladores, de tal manera que el Operador pueda controlar el proceso en modo Manual o Automático.

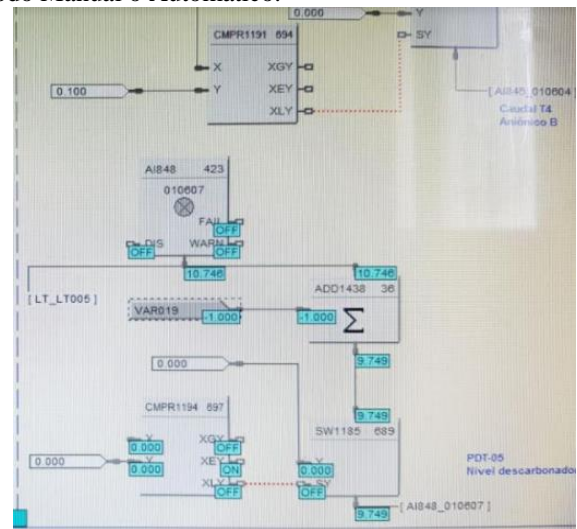


Figura 6. Estrategia del sistema Delta V para el control de nivel del tanque desaireador de la caldera
Fuente: Emerson – Delta V

Con el control manual, se manipula desde la pantalla la apertura y cierre de la válvula de paso de vapor/agua, en este caso se deja de lado la lectura del transmisor.

En el modo automático, todo el lazo de control trabaja y el operador únicamente visualiza los valores de pantalla y la válvula se abre o cierra dependiendo de los valores del transmisor.

IV. RESULTADOS EXPERIMENTALES

Una vez que se vio la necesidad de modificar el sistema de control por uno nuevo, que logre habilitar todos los procesos para la caldera en modo automático, se implementaron algunos cambios, entre ellos, el cambio de los transmisores y controladores, conjuntamente con un nuevo HMI, que controle el proceso.

Este cambio no fue fácil, ya que se tuvo que cambiar la mentalidad de los operadores y supervisores ya que el temor al cambio siempre provoca incertidumbre, por la falta de experticia en el nuevo proceso.

Con este nuevo sistema se ha logrado el control automático del 90% del proceso, con un aumento de confiabilidad en el control ya que se han adquirido las siguientes mejoras:

- 1.- Creación del historial de variables, lo que permite verificar los hechos pasados y mejorar los procesos siguientes, se puede evidenciar fallas de control y de instrumentos de campo que pueden ser mejorados.
- 2.- Implementación de nuevos controles para mejorar la eficiencia y seguridad de la Planta, a través de la información que se genera con los equipos directamente relacionados con la caldera.
- 3.- Actualizaciones del sistema y del Hardware para mantener los equipos con permanencia de seguridad y respaldo.
- 4.- Recuperación económica de la inversión, se estima que en 2 años se recuperó la inversión y fue un ejemplo para otras Plantas del mismo tipo.

Con el sistema anterior se tenía una confiabilidad del 60%, en cuanto tiene que ver con el sistema de control. Con el sistema Delta V, se tiene una confiabilidad de 90%.

En este contexto, se puede decir que los resultados obtenidos en este proyecto coinciden, en cuanto al aumento de la eficiencia, con otros proyectos relacionados con la implementación de sistemas de control distribuido en procesos productivos, así lo corroboran las investigaciones realizadas por Pérez- López [10]; Inserny, Aboul, Cañas y Custodio [11]; Manassero, Torres, López, Furlani, Regalini y Orué [12]; entre otros.

V. CONCLUSIONES

1. La implementación del sistema de control distribuido Delta V ha mejorado las condiciones operativas del caldero, haciendo el proceso más confiable, con herramientas para diagnóstico del funcionamiento de los equipos, lo que deriva en la optimización de recursos como el vapor de agua y el combustible.
2. El sistema de control Delta V es dinámico y va evolucionando a la par con la tecnología, por lo que se hace necesario actualizar el sistema conforme las recomendaciones técnicas del fabricante (Emerson), con esto se asegura que tanto el hardware como el software proporcionen las condiciones necesarias para mantener la confiabilidad del proceso.
3. Los altos costos que involucran la implementación del sistema Delta V, están justificados en plantas industriales

grandes, en las que los volúmenes de producción permiten la recuperación de la inversión en un tiempo adecuado.

VI REFERENCIAS

- [1] G. Correa-Henao y O. J. Pereira, Fundamentos de Control Automático E Inteligente, Omniscryptum GmbH & Company Kg, 2015.
- [2] V. Zárate Silva y P. Nájera García, Lo esencial de la instrumentación moderna para mecatrónicos, México: Editorial Digital del Tecnológico de Monterrey, 2013.
- [3] Creus, Instrumentación Industrial, España: Marcombo, 2012.
- [4] «PlantWEBUniversity,» Emerson Process Management, 2002. [En línea]. Available: http://www2.emersonprocess.com/siteadmincenter/PM%20Central%20Web%20Documents/EngSch-Buses_201. [Último acceso: 25 noviembre 2017].
- [5] Micro Motion Modbus Interface tool, Emerson Electric Co., St. Louis, Missouri, 2016.
- [6] N. Mahalik, Tecnología Fieldbus: estándares de red industrial para control distribuido en tiempo real, India: Springer Science & Business Media, 2013.
- [7] Rodríguez Penin, Sistema SCADA, Barcelona: Marcombo, 2012.
- [8] Manual de Ingenieros de Instrumentos, Volumen 3: Software de Procesos y Redes Digitales, Boca Ratón, FL., 2016.
- [9] Emerson Process Management, DELTAV DIGITAL AUTOMATION SYSTEM, Round Rock, TX, 2013.
- [10] E. Pérez-López, «Los sistemas SCADA en la automatización industria,» Revista Tecnología en Marcha, vol. 28, n° 4, pp. 3-14, 2015.
- [11] V. Inserny, A. Aboul, J. Cañas y Á. Custodio, «Diseño de un sistema de supervisión centralizado para las plantas de tratamiento de agua de CVG GOSH. Ciudad Guayana.» Universidad, Ciencia y Tecnología, vol. 12, n° 46, pp. 43-48, 2008.
- [12] U. Manassero, J. Torres, D. López, R. Furlani, R. Regalini y M. Orué, «Instrumentación, control y telesupervisión en centrales térmicas de pequeña potencia,» Iberoamerican Journal of Industrial Engineering., vol. 3, n° 2, pp. 70-86, 2011.

Análisis comparativo de metodologías de interpolación de polinomios que representen contornos de mapas para cálculo de áreas planas.

Wilson Benavides¹, Wilson Bravo¹
 wbenavides048@puce.edu.ec, wbravo670@puce.edu.ec
¹Pontificia Universidad Católica del Ecuador

Resumen— Englobar las diferentes aplicaciones en las ciencias y la ingeniería para obtener resultados importantes, requiere tener una visión amplia de la Matemática. Por ejemplo, en el caso de superficies planas delimitadas por contornos irregulares cuyas áreas principalmente se determinan por aproximaciones geométricas, el método más utilizado es el de Gauss en el que se aplica obviamente conocimientos matemáticos sin embargo podría obtener resultados más exactos si se aplican otros métodos basados en cálculos matemáticos que determinen el área con mayor exactitud. En esta propuesta se considera la medición del mayor número de puntos de un contorno, cuya representación son pares ordenados respecto a un plano cartesiano previamente escogido y construir mediante métodos de interpolación numéricos, polinomios para un intervalo de ese contorno irregular para posteriormente determinar las áreas hasta completar el contorno del mapa bajo los polinomios encontrados utilizando técnicas de integración mediante software especializado como el matemático Matlab o Scientific Workplace. La investigación consiste en comparar resultados obtenidos con diferentes metodologías de interpolación de polinomios que representan contornos de mapas para calcular su área mediante la integración definida entre las curvas logradas, como la de los polinomios interpolantes de Lagrange, Newton y Splines utilizando coordenadas cartesianas obtenidas por medición de puntos de su contorno. Es de particular interés perfeccionar varios temas de estudio de la Matemática que resalten y aporten en la formación de una carrera, con la relevancia que la misma representa en la visión y contextualización de su aplicación en diversos campos de las ciencias. La propuesta cobra importancia al ofrecer una metodología alternativa para el cálculo de áreas en zonas delimitadas por curvas irregulares en mapas, que involucre conocimientos de cálculo integral, *métodos numéricos* y manejo de software y que redundará en el análisis de la pertinencia de temas que se estudian en la Matemática Aplicada.

Palabras clave: Métodos de interpolación numéricos, polinomios interpolantes, polinomios interpolante de Newton, Polinomios interpolante de Lagrange, polinomio interpolador cúbico natural, integración numérica, software matemático

V. INTRODUCCIÓN

El papel de los métodos numéricos fue muy diferente en distintas áreas de la actividad humana y en distintas épocas. Se formó, históricamente, bajo la influencia considerable de dos factores: el nivel de desarrollo del aparato matemático sin computadora y luego con computadora.

Mediante el análisis numérico. Se construyen algoritmos que permiten calcular soluciones numéricas, generalmente asistidas

por computadora y las soluciones que proporcionan son numéricas y no en forma analítica. El modo numérico se enfoca sobre el proceso de computación y esto se debe establecerse explícitamente para aprovechar la potencia de la computadora.

En lo que diferencia el modo numérico es que se hacen a un lado por completo las operaciones analíticas y se usa desde el principio un planteamiento puramente numérico.

Respecto al problema del proyecto, las soluciones analíticas son complejas y por tanto debemos proceder a su aproximación.

Las soluciones aproximadas se calculan mediante la utilización de métodos y técnicas de interpolación e integración numérica adecuadas que permitan desarrollar algoritmos cuyos errores sean mínimos.

Interpolación

Interpolación significa estimar el valor desconocido de una función en un punto, tomando una media ponderada de sus valores conocidos en puntos cercanos al dado.

El problema de interpolación polinomial es el siguiente.

Dados $n + 1$ puntos en el plano (sin que haya dos en la misma recta vertical), el polinomio interpolador es el único polinomio de grado menor o igual que n que pasa por dichos puntos.

Este es el caso de un conjunto de datos calculados con una cierta precisión.

Varios son los métodos que se pueden aplicar para construir el polinomio de interpolación: usar polinomios de Lagrange, construir una tabla de diferencias divididas para emplearla con los coeficientes del polinomio de Newton o construir cúbicas con ciertos intervalos para los splines naturales.

Interpolación de Newton

En ocasiones resulta mejor construir polinomios de aproximación $P_1(x), P_2(x), P_3(x), \dots, P_n(x)$ y, después, elegir el más adecuado a nuestras necesidades. Si utilizamos los polinomios interpolantes de Lagrange, uno de los

inconvenientes es que no hay relación entre la construcción de mejor $P_{n-1}(x)$ y la de $P_n(x)$; cada polinomio debe construirse por separado y el trabajo necesario para calcular polinomios de grado elevado requiere hacer muchas operaciones.

Siguiendo un camino de construcción diferente, en el cual los polinomios interpolantes, que se denominarán de Newton, se obtienen mediante un esquema recursivo partiendo de un polinomio de primer orden (1), el de segundo orden(2) y así sucesivamente hasta llegar a uno de orden n (3)

$$P_1(x) = a_0 + a_1(x - x_0), \quad (1)$$

$$P_2(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1), \quad (2)$$

·
·
·
·
·

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + a_3(x - x_0)(x - x_1)(x - x_2) + a_4(x - x_0)(x - x_1)(x - x_2)(x - x_3) + \dots + a_n(x - x_0)(x - x_1)(x - x_2)(x - x_3) \dots (x - x_{n-1}). \quad (3)$$

Luego, $P_n(x)$ se reescribe en términos de $P_{n-1}(x)$

$$P_n(x) = P_{n-1}(x) + a_n(x - x_0)(x - x_1)(x - x_2)(x - x_3) \dots (x - x_{n-1}). \quad (4)$$

De esta manera decimos que el polinomio $P_n(x)$ es un polinomio de Newton con n centros $x_0, x_1, x_2, \dots, x_{n-1}$.

Como $P_n(x)$ involucra sumas de productos de factores lineales, con $(x - x_0)(x - x_1)(x - x_2)(x - x_3) \dots (x - x_{n-1})$ el de mayor grado, es inmediato darse cuenta que $P_n(x)$ es un polinomio de grado menor o igual que n .

Si se quiere obtener los coeficientes a_k de los polinomios $P_1(x), P_2(x), P_3(x), \dots, P_n(x)$ que nos sirven para aproximar una función dada $f(x)$. Entonces cada $P_k(x)$ es el polinomio de Newton que tiene como centros los puntos $x_0, x_1, x_2, \dots, x_k$ y es también el polinomio de interpolación para los nodos $x_0, x_1, x_2, \dots, x_{k+1}$.

Para el polinomio $P_1(x)$, los coeficientes a_0 y a_1 , indican que

$$P_1(x_0) = f(x_0) \quad \text{y} \quad P_1(x_1) = f(x_1) \quad (5)$$

Aplicando $P_1(x)$ para $x = x_0$ y $x = x_1$, se tiene

$$f(x_0) = P_1(x_0) = a_0 + a_1(x_0 - x_0) = a_0, \quad (6)$$

$$f(x_1) = P_1(x_1) = a_0 + a_1(x_1 - x_0) = f(x_0) + a_1(x_1 - x_0) \quad (7)$$

Despejando a_1 , obtenemos

$$a_1 = \frac{f(x_1) - f(x_0)}{x_1 - x_0} \quad (8)$$

El coeficiente a_1 es la pendiente de la recta que pasa por los puntos $(x_0, f(x_0))$ y $(x_1, f(x_1))$.

y así sucesivamente.

Es claro que los coeficientes a_k de los polinomios $P_n(x)$ dependen de los valores de interpolación $f(x_j)$, $\forall j = \overline{0, n}$.

Teorema del polinomio interpolador de Newton

Considere que $x_1, x_2, x_3, \dots, x_n$ son $n + 1$ números diferentes en $[a, b]$. Entonces existe un único polinomio $P_n(x)$ de grado menor o igual que n tal que

$$f(x_j) = P_n(x_j), \quad \forall j = \overline{0, n}. \quad (9)$$

La forma de Newton de este polinomio interpolador es:

$$P_n(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)(x - x_1) + \dots + a_n(x - x_0)(x - x_1) \dots (x_2 - x_{n-1}), \quad (10)$$

siendo $a_k = f[x_0, x_1, \dots, x_n]$, $\forall j = \overline{0, n}$.

Al aproximar una función $f(x)$ mediante el polinomio interpolador de Newton, la función toma la forma:

$$f(x) = P_n(x) + E_n(x), \quad (11)$$

donde $E_n(x)$ es el error de la función.

Interpolador de Lagrange

En la interpolación lineal se utiliza un segmento rectilíneo que pasa por dos puntos (x_0, y_0) y (x_1, y_1) conocidos, la pendiente de la recta viene dada por $m = ((y_1 - y_0)/(x_1 - x_0))$.

Esta podemos sustituir en la ecuación de la recta $y = m(x - x_0) + y_0$ y obtener

$$y = P(x) = y_0 + (y_1 - y_0) \frac{x - x_0}{x_1 - x_0} \quad (12)$$

El matemático francés Joseph Louis Lagrange descubrió que se puede encontrar este polinomio usando un método ligeramente distinto. Escribiendo

$$y = P_1(x) = y_0 \frac{x - x_1}{x_0 - x_1} + y_1 \frac{x - x_0}{x_0 - x_1} \quad (13)$$

Denotando los cocientes de la relación por

$$L_{1,0}(x) = \frac{x - x_1}{x_0 - x_1} \quad \text{y} \quad L_{1,1}(x) = \frac{x - x_0}{x_1 - x_0} \quad (14)$$

Teorema (Polinomio interpolador de Lagrange)

Sea $f \in C^{n+1}[a, b]$ y que $x_0, x_1, \dots, x_n \in [a, b]$ son $n + 1$ nodos de interpolación. Si $x \in [a, b]$, entonces $f(x) = P_n(x) + E_n(x)$, donde $P_n(x)$ es un polinomio que podemos usar para aproximar $f(x)$:

$$f(x) \approx P_n(x) = \sum_{k=0}^n f(x_k) L_{n,k}(x) \quad (15)$$

denominado polinomio interpolador de Lagrange de f para los nodos dados, y el error $E_n(x)$ se puede escribir como

$$E_n(x) = \frac{(x - x_0)(x - x_1) \dots (x - x_n) f^{(n+1)}(c)}{(n+1)!} \quad (16)$$

para algún valor $c = c(x)$ del intervalo $[a, b]$.

Interpolador cúbico natural (spline)

Una forma para ajustar un polinomio a un conjunto de datos $\{(x_0, f(x_0)), (x_1, f(x_1)), (x_2, f(x_2)), \dots, (x_n, f(x_n))\}$ es un trazador cúbico, denominado spline.

Un spline cúbico ajusta una "curva suave" a los puntos, tomando la idea de un instrumento que se usa para dibujar. Este instrumento es una barra flexible, que se flexiona para conformarse a los puntos (y que suele mantenerse en su sitio por medio de unos pesos). La curva spline puede ser de varios grados.

La ecuación para un polinomio cúbico, $g_i(x_i)$, en el i -ésimo intervalo, entre los puntos (x_i, y_i) y (x_{i+1}, y_{i+1}) es

$$g_i(x_i) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i \quad (17)$$

Así, la función del spline cúbico que se quiere, tiene la forma

$g(x) = g_i(x)$ sobre el intervalo $[x_i, x_{i+1}]$, $\forall i = \overline{0, n-1}$ y cumple las condiciones:

$$i) g_i(x_i) = y_i, \forall i = \overline{0, n-1} \text{ y } g_{n-1}(x_n) = y_n;$$

$$ii) g_i(x_{i+1}) = g_{i+1}(x_{i+1}), \forall i = \overline{0, n-2};$$

$$iii) g'_i(x_{i+1}) = g'_{i+1}(x_{i+1}), \forall i = \overline{0, n-2};$$

$$iv) g''_i(x_{i+1}) = g''_{i+1}(x_{i+1}), \forall i = \overline{0, n-2}.$$

Las ecuaciones i) indican que el spline cúbico se ajusta a cada uno de los puntos, que es continua ii) y que es continua y curvatura iii) y iv), a lo largo de toda la región generada por los puntos.

Si hay $n + 1$ puntos, el número de intervalos y el número de $g(x)$ es n . Así, hay cuatro veces n incógnitas, que son las $\{a_i, b_i, c_i, d_i\}$, $\forall i = \overline{0, n-1}$. Entonces se tiene la ecuación.

$$y_{i+1} = a_i h_i^3 + b_i h_i^2 + c_i h_i + y_i \quad (18)$$

Donde $h_i = (x_{i+1} - x_i)$, es el ancho del i -ésimo intervalo. Para relacionar las pendientes y las curvaturas de los splines de unión, se deriva la función $g_i(x_i) = a_i h_i^3 + b_i h_i^2 + c_i h_i + y_i$

El desarrollo se reduce si las ecuaciones se escriben en términos de las segundas derivadas, es decir, si se hace $S_i = g''_i(x_i)$, $\forall i = \overline{0, n-1}$

$$(19)$$

Despejando a_i y b_i , se tiene:

$$b_i = \frac{S_i}{2} \text{ y } a_i = \frac{S_{i+1} - S_i}{6h_i} \quad (20)$$

Usando la condición de las pendientes iguales de las dos cúbicas que se unen en (x_i, y_i) , la ecuación $g'_i(x_{i+1}) = g'_{i+1}(x_{i+1})$ en el i -ésimo intervalo con $x = x_i$ y de realizar las simplificaciones respectivas se obtiene

$$h_{i-1} S_{i-1} + 2(h_{i-1} + h_i) S_i + h_i S_{i+1} = 6(f[x_i, x_{i+1}] - f[x_{i-1}, x_i]) \quad (21)$$

VI. DESARROLLO

Metodología

La metodología empleada consistió principalmente en el desarrollo de 3 etapas.

1. Clasificación de mapas objetos de la investigación y cálculo de áreas por el método geométrico.

Esta fase consistirá en clasificar mapas que tengan perímetros irregulares, medición de puntos respecto a un sistema coordenado previamente seleccionado, obtención de bibliografía pertinente, análisis y elección de software adecuado para la investigación.

2. Elaboración de una tabla de valores para el cálculo de áreas de mapas utilizando el enésimo polinomio interpolante de Lagrange, diferencias divididas interpolantes de Newton, interpolación de trazadores cúbicos mediante coordenadas cartesianas obtenidas por medición de puntos de su contorno respecto a un sistema coordenado arbitrario que representan el contorno del mapa.

Se procederá a seleccionar el número de puntos a utilizar para establecer el dominio de los polinomios interpolantes a determinar y su posterior cálculo de áreas que encierran dichos polinomios, utilizando software matemático especializado como el Scientific Workplace.

3. Análisis y comparación de resultados entre la metodología implementada mediante los tres polinomios obtenidos para el mismo dominio y elaboración de un cuadro comparativo de los resultados.

VII. DISEÑO

Se consideró una región específica como área de estudio figura N°1 y fue la correspondiente a la provincia de Imbabura.

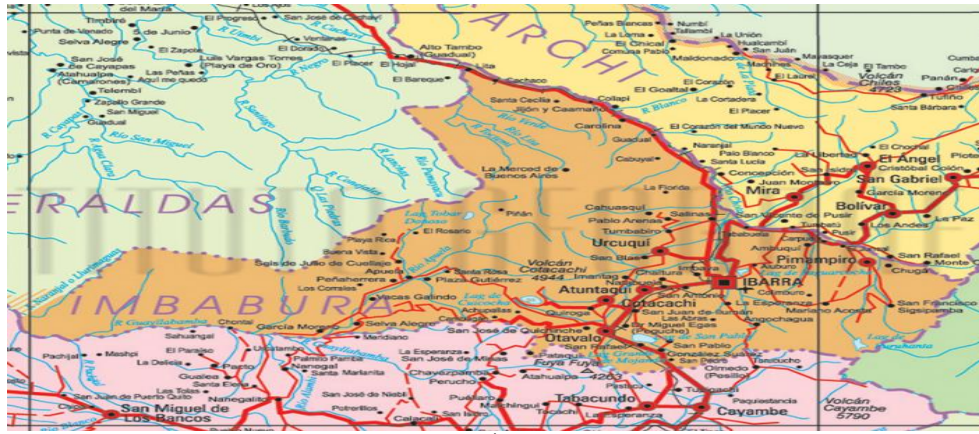


Figura N°1. Área de estudio

En el proyecto se discretiza el dominio y se consideran polinomios de cuarto grado para el método de Lagrange y de Newton por lo que se considerará 5 puntos por cuanto para obtener un polinomio de grado n se necesita $n+1$ puntos, mientras que para el polinomio por trazadores cúbicos se necesita 4 puntos.

Una vez obtenido los diferentes polinomios para cada método y para cada intervalo, se procede a unir para generar el contorno del mapa para el dominio de la región seleccionada.

Para el cálculo del área total se procede de forma similar que la anterior y se determina las áreas parciales bajo

cada polinomio obtenido hasta completar el área para todo el dominio de la región de estudio.

El software utilizado no es una herramienta de programación como Matlab, sin embargo, realiza una gran variedad de cálculos científicos y además es un editor de texto muy versátil para generación de artículos.

Los polinomios interpolantes de Lagrange encontrados para cada tramo tienen un color diferente y luego de unir varios intervalos resulta la parte superior del área de estudio figura N°2.



Figura N°2. Polinomios interpolantes de Lagrange

Uno de los polinomios de cuarto grado de Lagrange obtenido para un intervalo de puntos antes de unir, se muestra en la figura N°3

$$L_4(x) = 0,1152380952380952381x^4 - 6,54719047619047619x^3 + 138,648333333333333x^2 - 1297,6130952380952381x + 4551,705714285 \quad (22)$$

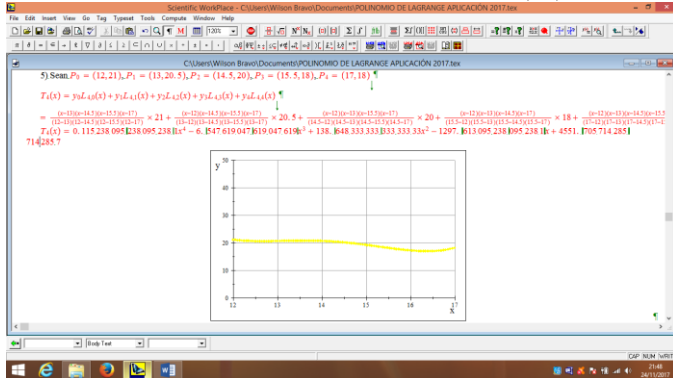


Figura N°3. Polinomio de Lagrange para un intervalo de puntos

Para otro intervalo de puntos, un polinomio de cuarto grado de Newton se muestra en la figura figura N°4

$$N_4(x) = 0,4888888888888888888x^4 - 84,2222222222x^3 + 5438,0555555555x^2 - 155974,3333333333x + 1676827,4 \quad (23)$$

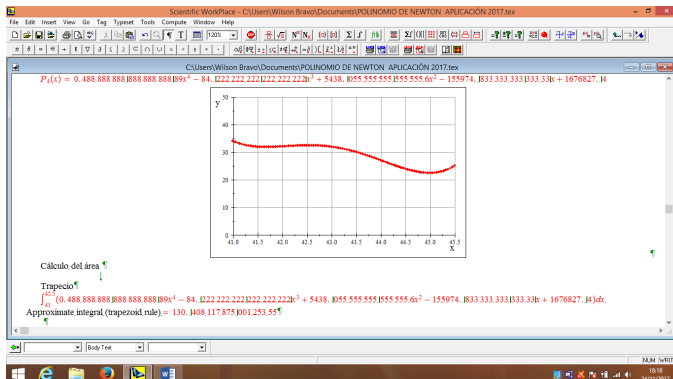


Figura N°4. Polinomio de Newton para un intervalo de puntos.

Una vez que se tienen los polinomios se procede a determinar el área bajo la curva mediante integración numérica empleando el método de los trapezoides y de Simpson.

El polinomio interpolante de trazador cúbico natural o splin natural para un intervalo de puntos se muestra en la figura N°5

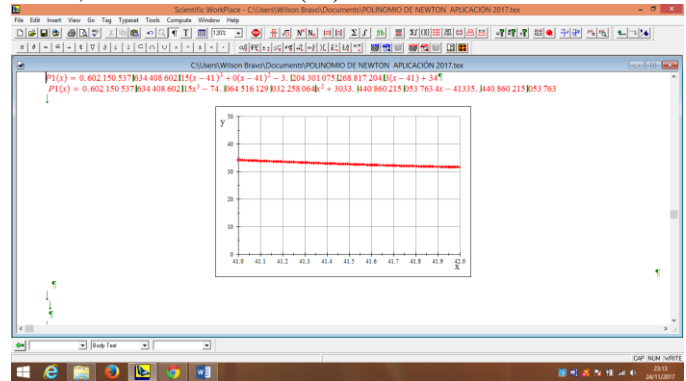
$$S_3(x) = 0,602150537x^3 - 74,064516129x^2 + 3033,440860215x - 41335,440860215 \quad (24)$$


Figura N°5. Polinomio interpolante de trazador cúbico natural o splin cúbico natural para un intervalo de puntos

Entre las características más sobresalientes de los polinomios obtenidos resalta la de que las pendientes entre punto y punto para su generación son constantes por lo que podría introducir errores por su aproximación en su cálculo, mientras que el polinomio interpolante obtenido por trazadores cúbicos o splines que se presenta en la figura N°5, se puede apreciar que entre punto y punto se generan pendientes variables por cuanto se unen mediante curvas de tercer orden y podría obtenerse una curva más exacta que represente de mejor manera el contorno del mapa analizado.

VIII. RESULTADOS EXPERIMENTALES

En la figura N°6 se muestra el gráfico de la superposición de los tres polinomios obtenidos para un mismo intervalo tramos rojo, azul y verde corresponden a spline, mientras que el turquesa a Lagrange y Newton y se puede apreciar las diferencias que tienen para el mismo intervalo de generación.

Los resultados de la integración numérica para el polinomio de Lagrange, Newton y de Trazadores cúbico o de splines se presentan en los cuadros de las figuras N°7, N°8, N°9.

V. CONCLUSIONES

1. La interpolación pueden ser representada mediante curvas que más se ajusten a los puntos experimentales obtenidos, representando de mejor manera las tendencias, siendo una alternativa que con menor error comparada con la interpolación lineal que es de uso muy frecuente y una puntos mediante líneas rectas.

2. Dentro de las curvas que más se ajustan a los puntos experimentales existen unas que se ajustan de mejor manera que otra a los puntos experimentales, en los métodos de Lagrange y Newton existe mucha variación en los extremos del intervalo, mientras que en el de splin cúbico natural no se tiene esa dificultad al momento de unir los gráficos de cada intervalo, análisis comparativo que es el objetivo de la presente investigación.

3. Se puede apreciar además diferencias entre las áreas calculadas bajo los tres polinomios interpolantes correspondientes a un mismo intervalo y calculadas por el método de trapecios y de Simpson, siendo valores muy parecidos entre el de Lagrange y Newton, sus gráficos se superponen de forma casi perfecta, por lo que el error en los dos métodos es igual.

4. El presente trabajo también se extiende puntos más exactos obtenidos mediante GPS, cuya latitud y longitud correspondiente se transforma a coordenadas cartesianas mediante procesos matemáticos de proyección a planos que generan los mapas de regiones.

5. Por las características de aproximación se puede concluir que el método de trazador cúbico natural o splin cúbico natural genera polinomios que se ajustan de mejor manera que el de Lagrange y Newton.

6. Por las características de aproximación se puede concluir que el cálculo de áreas por el método de Simpson genera menor error porque trabaja con trapecios curvilíneos a diferencia que el de trapecios que toma en consideración trapecios rectilíneos.

VI. REFERENCIAS

- [14] Raffo Lecca, Eduardo, Rojas Lazo, Oswaldo, Rojas Rojas, Jorge. "Cálculo de áreas de superficies irregulares aplicando software Matlab y Autocad". Lima: Revista de investigación UNMSM, (2007)
- [15] Atkinson, K.E., "An Introduction to Numerical Analysis", John Wiley, New York (1978).
- [16] Bajválov, N.S., "Métodos numéricos", Paraninfo, Madrid, (1966).
- [17] Bakhavalov, N., "Numerical Methods: Análisis, Álgebra, Ordinary Differential Equations", Mir, Moscú, 1977.
- [18] Cheney, W.; Kincaid, D., "Numerical Mathematics and Computing", Brooks/Cole, Monterrey (California) (1980).
- [19] Davis, P.J., "Interpolation and Approximation", Dover, New York (1975).
- [20] Rivlin, T.J., "An Introduction to the Approximation of Functions", Dover, New York (1981).
- [21] Stoer, J.; Bulirsch, R., "Introduction to Numerical Analysis", Springer-Verlag, New York (1980).
- [22] Tham, L. G., Li, W. Y. y Cheung, Y. K.: "Bending of skew plates by spline-finite-strip method", Computers & Structures, Vol. 22, No. 1, pp. 31-38, (1986).
- [23] Leung, A. Y. T. y Au, F. T.: "Spline finite element for beam and plate". Computers & Structures, Vol. 37, No. 5, pp. 717-729, (1990).

- [24] Luah, M. H. y Fan, S. C: "Analysis of complicated plates by a nine-node spline plate element", Computers & Structures, Vol. 45, No. 2, pp. 243-252, (1992).
- [25] Boor, C. y Lynch, R. E.: "On splines and their minimum properties". Journal of Mathematics and Mechanics, Vol. 15, No. 6, pp. 953-969, (1966).
- [26] Romero, J. L., Morant, J. L. y Corrales, J. M.: "Operadores de relajación en espacios de Hermite H (A). Aplicación al cálculo de splines", Congreso de Métodos Numéricos en Ingeniería, SEMNI, pp. 228-232, Las Palmas, (1990).
- [27] Romero, J. L., García, F. y Morant, J. L.: "Calculation of splines by iterative methods based on relaxation", en Proceeding of International Symposium of Numerical Analysis (ISNA '92), pp. 229-245, Prague, Czechoslovakia, (1992).
- [28] Romero, J. L., Corrales, J. M. y García, F.: "Interpolación mediante splines generalizados del tipo $a(x)D^n$ - splines", Congreso de Métodos Numéricos en Ingeniería, SEMNI, pp.720-730, La Coruña, (1993).

La Simulación en Ingeniería

Transcendiendo Fronteras



Pontificia Universidad
Católica del Ecuador

ISBN: 978-9942-9902-6-6

