



**UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS**
FACULTAD DE INGENIERÍA

Revista Ingeniería

Volumen 18 • Número 1 • Año 2013 • ISSN 0121-150X

REVISTA CIENTÍFICA SEMESTRAL

2013

Revista Ingeniería

Volumen 18 • Número 1 • Año 2013 • ISSN 0121-150X • E-ISSN 2344-8393

REVISTA CIENTÍFICA SEMESTRAL

Publicación admitida en el Índice Nacional de Publicaciones Seriadadas Científicas y Tecnológicas de Colciencias



UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS
FACULTAD DE INGENIERÍA

Carrera 7 No. 40-53 Piso 4 Bogotá, Colombia
Teléfono: + 57 (1) 323 93 00 ext. 2413
Correo electrónico: revista_ing@udistrital.edu.co
<http://revistas.udistrital.edu.co/ojs/index.php/reving>

EDITOR

Sergio A. Rojas, PhD.

Universidad Distrital Francisco José de Caldas

COMITÉ EDITORIAL

Luz Esperanza Bohorquez, MSc.

Universidad Distrital Francisco José de Caldas

Marco Alzate, PhD.

Universidad Distrital Francisco José de Caldas

Henry Diosa, PhD.

Universidad Distrital Francisco José de Caldas

Rodrigo Herrera, MSc.

Universidad Distrital Francisco José de Caldas

Ana María Peña, PhD.

Universidad Distrital Francisco José de Caldas

Edwin Rivas, PhD.

Universidad Distrital Francisco José de Caldas

Juan Carlos Figueroa, MSc.

Universidad Distrital Francisco José de Caldas

COMITÉ CIENTÍFICO

Germán Méndez, PhD.

Universidad Distrital Francisco José de Caldas

Víctor Hugo Grisales, PhD.

Universidad Nacional de Colombia

DIRECTIVAS

Inocencio Bahamón C., MSc.

Rector

María Elvira Rodríguez, PhD.

Vicerectora Académica

Octavio Salcedo Parra, MSc.

Decano Facultad Ingeniería

Ing. Felipe Forero, CAE (Cambridge)

SCIENTIFIC WRITING ADVISOR

Diagramación y carátula

Guillermo Peñalosa Martínez

IMPRESIÓN

Editorial UD

Universidad Distrital Francisco José de Caldas
Carrera 19 No. 33 -39.

ÁRBITROS EN ESTA EDICIÓN

Rodrigo A. Gómez Montoya, PhD(c).

Politécnico Colombiano Jaime Isaza Cadavid

Oscar Herrera Ochoa, MSc.

Universidad Cooperativa
de Colombia

Willy H. Calsina Miramira, MSc.

Universidad San Martín de Porres

Elizabeth León Guzmán, PhD.

Universidad Nacional de Colombia

Juan Carlos Ángel, MSc.

CENICANA

Edwin Rivas, PhD.

Universidad Distrital Francisco
José de Caldas

Cristina Gonzales Caro, PhD.

Universidad Autónoma de Bucaramanga

Fidel Torres, PhD.

Universidad de los Andes

José Luis Villa, PhD.

Universidad Tecnológica de Bolívar

POSTURA EDITORIAL Y AUDIENCIA

La Revista INGENIERÍA es editada por la Universidad Distrital, y publicada con periodicidad semestral. Su principal objetivo es difundir y debatir avances en investigación y desarrollo en las diferentes áreas de la ingeniería. El contenido de la revista está dirigido a investigadores, especialistas, docentes, estudiantes de pregrado a posgrado, ingenieros en ejercicio y al sector empresarial.

PERMISO DE REPRODUCCIÓN

Los textos de los artículos incluidos en esta edición pueden ser utilizados y reproducidos con fines sin ánimo de lucro y dando crédito a los autores.



Este trabajo está bajo una Licencia de Creative Commons Reconocimiento – No comercial – Sin Obra Derivada 3.0 Unported

SUSCRIPCIÓN

Cra. 7 No. 40-53 Piso 4 Decanatura Facultad de Ingeniería
Teléfono: + 57 (1) 3239300 ext. 2413
<http://revistas.udistrital.edu.co/ojs/index.php/reving/user/register>

INDIZADA EN

Publindex
Colciencias

PERIÓDICA
UNAM

latindex
UNAM

Dialnet
Universidad de la Rioja

**Red Colombiana de
Revistas de Ingeniería**

**INDEX COPERNICUS
INTERNATIONAL**

Ni la Revista INGENIERÍA, Editor, Comité Editorial, Comité Científico, Facultad de Ingeniería de la Universidad Distrital F.J.C., otorgan ninguna garantía, expresa o implícita, a asumen responsabilidad alguna por la exactitud, completitud o utilidad de cualquier información, aparato, producto o proceso divulgado, o que represente que su uso no infrinja derechos privados. La mención o referencia a algún producto, proceso o servicio comercial en específico, por su marca comercial, marca registrada, fabricante o cualquier otra denominación, no implica ni constituye su endoso, recomendación o favorecimiento por parte de la Revista INGENIERÍA. Los juicios y opiniones expresadas por los autores en este medio son de su responsabilidad y no establecen, reflejan o comprometen los de la Revista INGENIERÍA.

FE DE ERRATA: La Revista Ingeniería presenta excusas al PhD. Henry Diosa, por la omisión del nombre como miembro del Comité Editorial en las ediciones v17n1 y v17n2, siendo un error de edición en los volúmenes correspondientes.

CONTENIDO

EDITORIAL

-
- Revistas Colombianas de Ingeniería: ¡Uníos!** 4
Sergio A. Rojas
-

ARTÍCULOS DE INVESTIGACIÓN

-
- Coloración Psicodélica de Imágenes Digitales mediante Lógica Difusa** 6
Psychedelic Coloring of Digital Images Through Fuzzy Logic
Linda Sheriyn Rodríguez Castro • Iván Javier Ponce Fajardo
Helbert Eduardo Espitia Cuchango
-

- Extensión de taxonomía y tratamiento de valores faltantes sobre un repositorio de incidentes de seguridad informática** 24
Taxonomy extension and missing-values treatment over an informatics-security incident repository
Carlos Javier Carvajal Montealegre • Diego Nicolás Bayona
Zulima Ortiz Bayona
-

ARTÍCULOS DE REVISIÓN

-
- Identificación de relaciones entre los nodos de una red social** 50
Identification of node relationships in a social network
Mónica Andrea Niño Barón • Sonia Ordóñez Salinas
-

- Ubicación y dimensionamiento como parámetros en el diseño de almacenes: revisión del estado de arte** 65
Location and sizing as parameters in the ware-house design: A survey paper
Andrés Mauricio Hualpa Zúñiga • Carolina Suárez Roldan
-

- Gestión de la cadena de abastecimiento del biodiésel: una revisión de la literatura** 84
Biodiesel supply chain management: A survey paper
Manuel Barón • Isaac Huertas • Javier Orjuela
-

- Una revisión al estado del arte de la integración de toma de decisiones en la red logística** 118
State of the art review of the integration of decision making in the logistics network
Cesar Augusto Pineda Pérez • Cesar Amílcar López Bello
-

Editorial

Revistas Colombianas de Ingeniería: ¡Uníos!

Tal como había sido anunciado, Colciencias se encuentra sensibilizando a la comunidad sobre la inminente puesta en marcha del nuevo modelo de indexación de publicaciones seriadas de ciencia tecnología e innovación – Publindex. Aunque la versión final difiere en algunos aspectos de la inicialmente planteada (ver nuestra nota editorial 2011-I, Vol.16, Num.1), el espíritu de la norma sigue siendo propender por aumentar los niveles de visibilidad e impacto de las publicaciones locales de acuerdo con los estándares de medición de índices bibliográficos internacionales (ISI, Scopus, y otras bases de datos bibliográficas). Los criterios más estrictos se refieren a aumentar la cota mínima de artículos de investigación (12 por año para C, 15 para B) y articular las categorías A1, A2, A3, A4 (aparecen estas dos últimas nuevas) a los cuartiles del índice de impacto de ISI o Scopus, y las categorías B y C a las bases de datos bibliográficas Scielo y Redalyc para el caso de ingeniería.

En nuestra opinión el espíritu de la normativa propuesta es noble en cuanto pretende fortalecer la dinámica de la divulgación científica en nuestro país, eliminar la endogamia y proyectar desde lo local hacia lo universal la transferencia de resultados de nuevo conocimiento e innovación tecnológica que logran nuestros investigadores. Sin embargo nos parece que solamente endurecer los criterios de indexación no es suficiente. Pues no encontramos detrás de esta propuesta una política de acompañamiento de parte de Colciencias, de forma que se definan estrategias y recursos que respalden la adopción del nuevo modelo y que permitan a las revistas nacionales equipararse con sus pares internacionales. En este sentido persisten las dudas que nos planteábamos en dicho editorial: ¿cómo garantizar que la transición a referentes universales sea continua y no abrupta? Ya que para ello se requieren gestión, conocimiento, apoyo técnico y financiero por fuera del alcance de la mayoría de nuestras revistas en su grado de madurez actual, poniendo en riesgo de fracaso los procesos que en mayor o menor grado, han recorrido hasta el momento las revistas. Por otra parte, ¿es viable medir efectivamente los alcances de impacto local y/o regional que puedan tener investigaciones sobre problemas inherentemente colombianos a través de índices internacionales, que por demás son manejados por compañías privadas con limitaciones de acceso libre para la comunidad? Creemos que no, mientras no se consolide una dinámica de reconocimiento y citación nacional de los trabajos publicados en nuestras revistas, que en la actualidad no ocurre debido a la proliferación que indirectamente causó el modelo Publindex anterior. Hoy en día tenemos prácticamente una o más revistas por cada facultad de ingeniería del país (62)¹, dispersando fortalezas y esfuerzos que de no unificarse, difícilmente podrán alcanzar la visibilidad requerida.

Finalmente, quizás la pregunta mas inquietante, ¿cuántas de nuestras revistas hoy en día cuentan con las estructuras de apoyo y madurez de procesos necesarias para alcanzar el nivel A del nuevo índice? Tres o cuatro es nuestra predicción, ubicándose a lo más en A4. Surge pues la inquietud: ¿Cuál es el horizonte avizorado

1 Número de revistas indexadas en Publindex en 2013, dentro de la gran área de conocimiento: ingenierías.

por Colciencias para que las demás lleguen allí, o más aún, para que estas pocas suban hasta A1? Lastimosamente, luce oscuro. A menos que exista un proceso de crecimiento continuo respaldado con acompañamiento institucional, gremial, comunitario y estatal, que permita alcanzar visibilidad nacional a corto plazo (Colombia) y regional a mediano plazo (Latinoamérica). Aparte del modelo de medición, ¿cómo va a apoyar Colciencias tal proceso? Desde la Revista INGENIERÍA queremos lanzar una propuesta que contribuya a iniciar un debate que quizás contribuya a resolver estas preguntas. Partiremos de la premisa de que el nuevo modelo, a la luz de los cuestionamientos expresados, desvirtúa la naturaleza de contar con revistas multidisciplinarias de ingeniería editadas por universidades o facultades con población objetivo local. La propuesta consiste pues, en buscar una fusión de ellas para dar origen a una «Revista Colombiana de Ingeniería» unificada, con secciones especializadas en las diversas ramas de la ingeniería que se practica en Colombia, con una periodicidad bimestral o mensual (según el número de especialidades), y editada con el apoyo de organizaciones científicas o gremiales nacionales. Con el ánimo de evitar posibles conflictos de interés o luchas de poder, la dirección editorial de esta nueva revista debería desprenderse de las universidades o instituciones editoriales actuales, para depositarla en un organismo imparcial encargado de realizar su gestión (proponemos por ejemplo a la Asociación Colombiana para el Avance de la Ciencia – ACAC o a la Sociedad Colombiana de Ingenieros). En cuanto a los recursos financieros necesarios para su funcionamiento, consideramos que deberían ser provistos conjuntamente por Colciencias (como parte de la estrategia que respalde el nuevo modelo de medición) y ACOFI (como representante de facultades de ingeniería que agrupa y de las cuales recibiría aportes para este fin). En cuanto a los recursos humanos, obviamente serían aportados por las universidades y entidades editoras actuales, conformando equipos altamente calificados de editores, comités científicos y árbitros. La nueva Revista debería tener al menos un representante de cada institución. Al respecto se requeriría igualmente el compromiso de las instituciones para reconocer la profesionalización de la labor de los editores y revisores, autorizando por ejemplo, horas en sus planes de trabajo para los profesionales que contribuyan en dichas tareas. Por último, para evitar demoras en la periodicidad debido a inconvenientes con la impresión, se propone que la nueva Revista sea electrónica, y que Publindex la indexe inicialmente con la moda de las categorías de aquellas que se fusionen, con una validez no menor a 2 años mientras logra alcanzar los criterios de existencia requeridos por el nuevo modelo, antes de ser re-categorizada.

Suponemos también que esta iniciativa favorecería en mayor grado a las revistas que en la actualidad nos encontramos en categorías bajas de Publindex, que a las que se han consolidado con arduo trabajo en las más altas. Pero por otro lado reconocemos que el número de revistas en categorías C y B es mucho mayor que las de A1 y A2, por lo que se podría conformar masa crítica y sobre todo aunar las fortalezas de editores y comités editoriales que en ellas están dispersas, para que unificadas posibiliten su pronta ubicación en los índices internacionales. Nótese que la Revista unificada tendría fácilmente visibilidad nacional inmediata, con lo cual se viabiliza el fortalecimiento de su factor de impacto, en el corto plazo nacional, y en el mediano plazo, regional. Una vez fortalecida, podría convocar a las actuales A1 y A2 para intentar convertirse en un órgano de difusión científico de impacto universal, como lo pretende el nuevo modelo de medición.

Por el momento continuamos aportando con la presente edición, artículos sobre: sistemas difusos para coloración psicodélica, taxonomías de incidentes de seguridad informática, gestión de la cadena productiva del biodiesel, identificación de entidades nombradas en redes sociales y análisis de parámetros de ubicación y dimensionamiento en almacenes. Agradecemos sinceramente como siempre a nuestros autores, revisores y lectores. Y a la comunidad académica la exhortamos a reflexionar en torno a la conveniencia de unificar desde ya fortalezas, experiencias y entusiasmos.

Ing. Sergio A. Rojas, MSc., PhD.
Editor General Revista INGENIERÍA

Coloración Psicodélica de Imágenes Digitales mediante Lógica Difusa

Psychedelic Coloring of Digital Images Through Fuzzy Logic

Resumen

En el presente artículo se propone un sistema de inferencia difusa para la manipulación del color en imágenes digitales con el propósito de modificar su estilo artístico. Se genera una combinación de colores diferente a la original de tal forma que produzca una nueva sensación visual. El sistema de lógica difusa propuesto, modela mediante conjuntos y reglas de inferencia difusas, la modificación de color considerando la modalidad artística denominada arte psicodélico, generando como resultado imágenes cuyos colores producen este efecto. En la segunda parte se analiza el sistema de inferencia propuesto según datos empíricos determinados junto con expertos en el campo del arte, con el fin de obtener un mejor ajuste a la escala de colores del estilo artístico estudiado. Mediante la presentación de diferentes casos de prueba se evidencian los resultados obtenidos por el sistema antes y después de la optimización.

Palabras clave: Arte evolutivo, colorimetría, lógica difusa, procesamiento de imágenes.

Abstract

This paper proposes a fuzzy inference system for color manipulation in digital images in order to modify image artistic style. This manipulation generates a different color scheme that differs from that of the original image and produces a new visual feeling. Color modification is achieved through a fuzzy-logic system composed of fuzzy sets and rules. The study considers an art form known as psychedelic art, generating (as a result) images whose colors cause similar effects. In the second part of this report, the system is optimized by using empirical data gathered from a panel of art experts in order to enhance the color map of the artistic style. Through different test cases, the results obtained by the system before and after optimization are shown.

Key words: Evolutionary art, colorimetry, fuzzy logic, image processing.

Linda Sheriyn Rodríguez Castro
Universidad Distrital
Francisco José de Caldas
Facultad de Ingeniería
lsrodriguez@correo.udistrital.edu.co

Iván Javier Ponce Fajardo
Universidad Distrital
Francisco José de Caldas
Facultad de Ingeniería
ijponcef@correo.udistrital.edu.co

Helbert Eduardo Espitia Cuchango
Universidad Distrital
Francisco José de Caldas
Facultad de Ingeniería
heespitac@udistrital.edu.co



Fecha recibido: may.05/2013
Fecha modificado: jun.06/2013
Fecha aceptado: jun.14/2013



1. Introducción

El arte evolutivo se ha convertido en una importante rama de investigación que establece un punto de encuentro entre el campo de la expresión humana propiamente dicho y disciplinas como las ciencias de la computación y el modelamiento de sistemas de control. El principal objetivo de los estudios clasificados dentro del arte evolutivo es lograr traducir los procesos humanos de creatividad en algoritmos computacionales [1] que generen obras artísticas y resultados verificables por expertos, o modelos automáticos de validación [2].

A través de la implementación de sistemas de tratamiento del color, se han obtenido resultados y aportes en la detección de estructuras específicas dentro de imágenes a través de sus características de color; como por ejemplo la selección apropiada de colores para subtítulos en películas a color [3]; la detección del fuego en cámaras de seguridad [4], el mejoramiento en las búsquedas web de imágenes reconociendo un objeto particular en la imagen por su color [5], la reducción de ruido en imágenes de color [6], la optimización de tareas de procesamiento multimedia cuantificando el color para reducir la complejidad mediante el uso de un menor número de bits para representar el color verdadero, basándose en la similitud de codificación de color en vecindades de bits, esto aplicado a la detección de cambio de escenas en videos [7]; sistemas de clasificación de color basados en segmentación de imágenes y optimizados con algoritmos inspirados en enjambres para reducir el error [8].

La lógica difusa permite una aproximación a la clasificación del color desde la perspectiva humana mediante técnicas que no establecen límites cerrados, generando cierto nivel de incertidumbre [9] en contraste con la dicotomía encontrada en la lógica computacional que maneja estándares absolutos que van de valores de 0 a 1 [10]. Este contraste se evidencia en la clasificación del color al establecer la taxonomía subyacente al estudio de un color particular, utilizando lógicas booleanas el valor de 1 o 0 afirmarían si el color pertenece a una denominación en particular que a su vez es de carácter excluyente, mientras la lógica difusa da la posibilidad de fijar reglas y conjuntos con amplios rangos de pertenencia que desdibujan las fronteras, generando sistemas donde términos como “verde rojizo” están totalmente permitidos mediante niveles de pertenencia a los diferentes grupos [11].

Los diferentes grupos dados por la lógica difusa en el proceso de percepción del color pueden apoyarse en la representación dada por el modelo RGB que permite ubicar un color a través de códigos que evalúan sus componentes o niveles de pertenencia a tres diferentes conjuntos: rojo, verde y azul [3].

La utilización de la lógica difusa para el manejo del color, es una de las muchas aplicaciones dentro del contexto del procesamiento de imágenes, ya que su fin en la mayoría de los estudios es el análisis de las características propias de una entrada para su posterior modificación atendiendo a ciertos requisitos. Los tres principales pasos de aplicación de lógica difusa en el procesamiento de imágenes son: fusificación de la imagen, establecimiento de un sistema de inferencia y defusificación.

En el campo del arte también se encuentran aplicaciones de la lógica difusa, aunque en menor medida. Entre éstas vale la pena mencionar la integración del lenguaje natural que provee información básica del contexto de paisaje para aplicar color de forma inmediata a un cuadro [12]; así como la estrategia de mapeo para la composición audiovisual y nuevas técnicas de síntesis de sonido [13] que se introducen en el campo de lo que hoy se denomina arte evolutivo.

El desarrollo propuesto en el presente artículo se relaciona con el procesamiento de imágenes, para la identificación de color y modificación de la expresión artística de éstas, es decir aquello que buscan transmitir a través de diferentes combinaciones de colores y técnicas. Ejemplificando dicho tratamiento visual se propone la generación de características asociadas con el arte psicodélico en imágenes digitales. El término psicodélico fue acuñado por el psicólogo británico Humphry Osmond, quién afirmaba que su significado refería “lo que manifiesta la mente” [14]. En el arte, específicamente en la pintura, esta manifestación se expresa a través de visiones del mundo no convencionales donde las características propias del uso del color y las formas juegan un papel realmente importante. Dentro de las características se encuentran el uso de patrones fractales, colores brillantes y con alto contraste, profundidad en los detalles, y formas de tipo radial, como espirales. Los detalles propios de la pintura psicodélica son adecuados para la presentación de resultados de la implementación presentada en este artículo, por la evidente transformación de los datos de entrada.

Con fundamento en el carácter propio del estilo artístico descrito y el modelo de inferencia difusa para la transformación de color se propone a continuación un sistema en el que las variables de entrada y salida, las reglas lógicas condicionales, las funciones de pertenencia y demás elementos conlleven a la modificación de los colores mediante la disminución de los rangos del modelo RGB para adquirir los efectos deseados. Este trabajo constituye un primer paso hacia sistemas de modificación visual no enmarcados dentro de un único estilo artístico.

La importancia del desarrollo propuesto radica en la generación de sistemas inteligentes con capacidad de toma de decisiones basadas en conocimiento. Esto se logra mediante la parametrización del conocimiento necesario para la transformación de imágenes en un determinado estilo artístico. Con una ayuda de expertos en un área especializada del conocimiento, como las artes, se logra automatizar una tarea que tradicionalmente requiere del criterio profesional. De esta forma, los resultados que se presentan en este artículo evidencian el acercamiento al propósito expresado, mediante ejemplos concretos de transformación del estilo artístico para diferentes tipos de imágenes.

Este documento se divide en seis secciones. En la Sección 1 encontramos la introducción. La Sección 2, Sistemas de Lógica Difusa, presenta de forma breve los conceptos asociados a los sistemas de lógica difusa necesarios para la comprensión del desarrollo propuesto. La Sección 3, Sistema de inferencia difusa para la modificación



del color, presenta la definición del sistema propuesto para la coloración psicodélica de imágenes digitales, subdividido en: explicación del problema, funciones de pertenencia y reglas. En la Sección 4, Ajuste de los conjuntos de pertenencia, se expone la optimización del sistema mediante el método del gradiente con el apoyo de expertos en el campo del arte. La Sección 5 evidencia los resultados obtenidos por el sistema antes y después de su optimización. Finalmente, la Sección 6 argumenta las conclusiones y expectativas del trabajo futuro a realizar.

2. Sistemas de Lógica Difusa

La Lógica difusa es una técnica de inteligencia computacional cuyos insumos provienen de múltiples variables, y que intenta controlar la incertidumbre proveniente de las características imprecisas dadas en el mundo real. Además de la incertidumbre, la ambigüedad y su manejo es la principal motivación de la lógica difusa. La vaguedad a su vez se relaciona con el conocimiento del valor de una función en términos de una variable con valor conocido [15].

Las funciones asociadas a los sistemas de inferencia difusa son conocidas como funciones de pertenencia, que representan las entradas y salidas en términos de variables que se relacionan mediante un conjunto de reglas que ayudan a obtener un modelo que simula el comportamiento propio de la situación de incertidumbre que se estudia. Estos elementos, en resumen, buscan establecer el nivel de pertenencia de un elemento a un conjunto, y a la vez disminuir la diferencia entre los valores encontrados en los sistemas a modelar y su representación difusa.

El desarrollo de un modelo de lógica difusa comienza por un proceso denominado fusificación, producto del cual se obtienen las variables de entrada y salida junto con las funciones de pertenencia que describen su comportamiento. Posteriormente, mediante la etapa de inferencia lógica, se establecen las reglas de implicación que llevan a los datos de entrada a las salidas determinadas. Finalmente, para obtener de nuevo valores que tengan significado para el sistema del mundo real, se lleva a cabo el proceso de defusificación.

En este trabajo se aplica la lógica difusa porque su enfoque permite que es bien sabido que es un enfoque que permite abordar problemas de un alto grado de complejidad, no lineales, que no cuentan con modelos matemáticos precisos que los representen, o cuyo conocimiento asociado sea de carácter subjetivo [15].

3. Sistema de inferencia difusa propuesto para la modificación del color

En esta sección se describen las etapas asociadas a la definición del sistema de lógica difusa propuesto para la coloración psicodélica de imágenes digitales.

3.1. Descripción del problema

Dada la ambigüedad en la selección de un color perteneciente al conjunto de colores del arte psicodélico y partiendo de colores de una imagen con diferente percepción artística, se busca alcanzar una sensación visual psicodélica. Para ello, se hace necesaria la definición de rangos de color que pueden ser descritos y especificados a través de obras digitales mundialmente conocidas, producidas por artistas dentro del estilo psicodélico.

A manera de ilustración la Figura 1 y la Figura 2 muestran ejemplos de arte psicodélico. Cómo puede detallarse los colores predominantes son diferentes tonalidades de azules, morados, fucsias, verdes y naranjas. Basándose en estas observaciones junto con el análisis visual de otra variedad de obras de artistas ampliamente conocidos dentro del estilo artístico psicodélico como Larry Carlson, Marta Minujín, Richard Avedon, Alex Grey, Gilbert Shelton, entre otros, se obtuvieron las reglas de inferencia que se describen a continuación.



Figura 1. Colorized Afternoon. (Tarde Coloreada).
Autor: Larry Carlson. (2009).

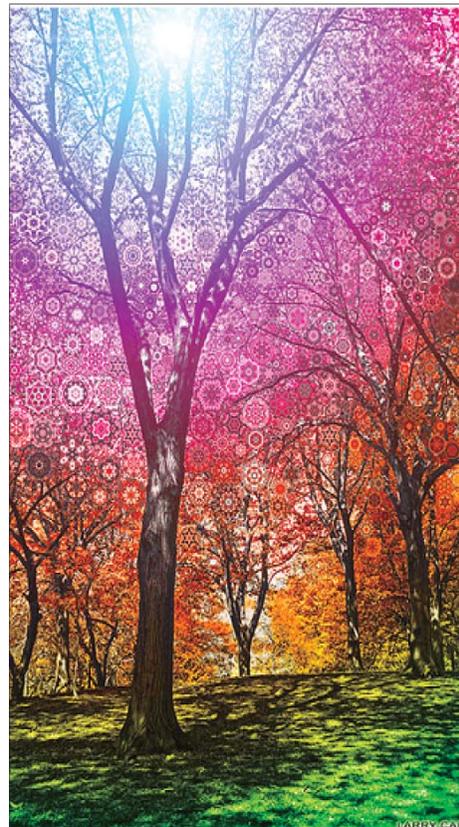


Figura 2. The Wonder Tree. (El Árbol Maravilloso).
Autor: Larry Carlson. (2011).



3.2. Funciones de pertenencia propuestas

El universo discurso de entrada del sistema de inferencia difusa está dado por los datos aportados por las imágenes que se modificarán, constituido por tres variables diferentes: una para el color rojo, otra para el color verde y otra más para el azul, correspondiendo así a la del modelo de color utilizado (modelos RGB – *Red Green Blue*). Estas tres variables utilizan rangos de valores de 0 a 255, que establecen el nivel de aporte del color correspondiente dado para un determinado pixel.

Para cada variable de entrada se utilizan tres funciones de pertenencia de tipo gaussiana que denotan tonos oscuros, medios y claros para el color determinado. Las distribución de las funciones de pertenencia para cada una de las variables de entrada se observa en la Figura 3 parte (a).

Los rangos para cada función de pertenencia asociadas a las variables de entrada (color original) se establecieron mediante el análisis visual del espectro de color para cada componente, específicamente para la componente roja se determinó un rango aproximado de 0 a 160 para tonos oscuros, 120 a 225 para tonos medios y 155 a 255 para tonos encendidos. El mismo proceso de análisis dio como resultado diferentes rangos para las componentes de color restantes: verde y azul, puesto que las tonalidades verdes tienden a percibirse más claras en un grupo más amplio de valores. A su vez, los tonos medios del color verde se orientan más hacia los tonos opacos. En particular para el color azul, los tonos medios se extienden a más valores al igual que las tonalidades oscuras.

El universo discurso de salida se compone por tres variables, que de igual forma que la entrada representan las componentes del modelo de color que permiten representar los nuevos colores pertenecientes a conjuntos con características del arte psicodélico.

Según el criterio de los expertos, los colores predominantes en este estilo artístico estudiado corresponden a tonos saturados, es decir, donde dos de las componentes del modelo RGB poseen valores bajos con respecto a la otra componente. Por esta razón se acota el rango de cada una de las funciones de pertenencia para las variables de salida.

Como se evidencia en la Figura 3 parte (b) las funciones de pertenencia para las variables de salida (color modificado – psicodélico) utilizadas son de tipo gaussiana, correspondiendo igualmente a tonos oscuros, medios y claros. La disposición de las variables abarca para cada una, un menor rango de valores con el fin de obtener cierta gama de colores más limitada y acotada por las características definidas.

3.3. Reglas propuestas

Una vez establecidos los rangos de colores comunes en el estilo visual que se desea proporcionar a las imágenes que ingresan en el sistema, se determina cuáles tonalidades serán reemplazadas y qué combinación de los diferentes conjuntos difusos llevarán a tal resultado. En la Tabla I se consignan la totalidad de las reglas establecidas.

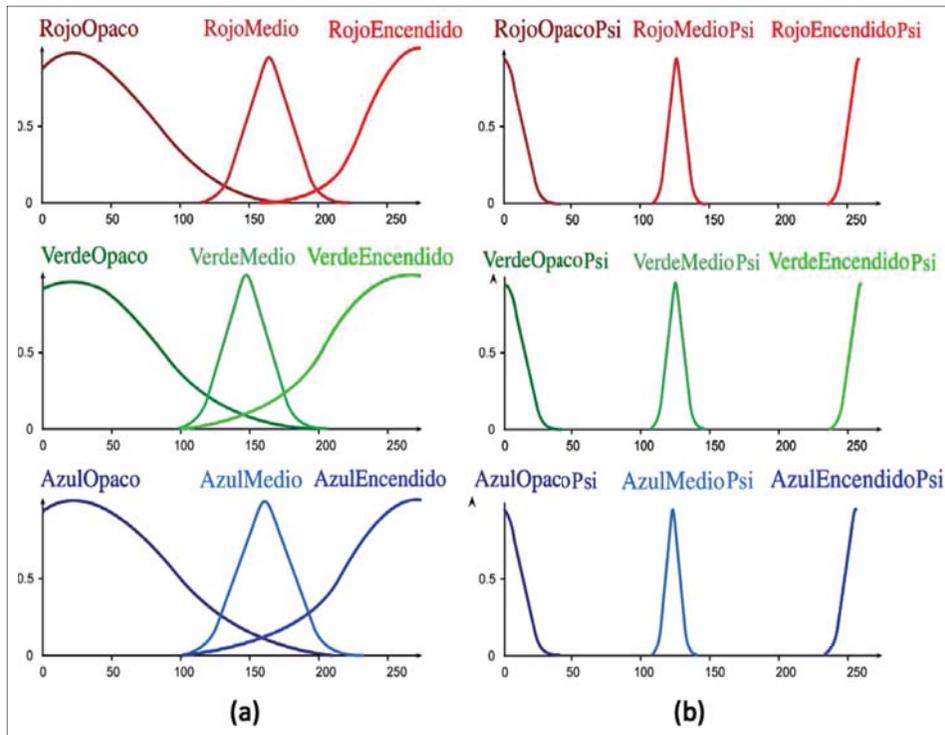


Figura 3. Funciones de pertenencia del sistema de inferencia.
 (a) Variables de Entrada (color original) (b) Variables de salida (color psicodélico)

A manera de ejemplo, la regla 10 define la siguiente transformación: si el pixel es rosado (rojo encendido + verde medio + azul medio) se colorea fucsia (rojo encendido + verde opaco + azul medio). La regla anterior y las otras 16 reglas se obtuvieron a partir del análisis experimental con grupos acotados de color e información teórica dada por expertos, como la obtención de un tono saturado o puro que enmarcan el estilo artístico estudiado.

La obtención de tonos saturados se consigue mediante el aumento en el valor de la componente de color con mayor número frente a las otras dos y como consecuencia la disminución de valor en esas dos componentes. En los casos de las reglas 1, 2 y 3 se caracterizan por llevar a la componente roja, verde y azul respectivamente a un región muy acotada con valores entre los 230 y los 255 aproximadamente y a las dos componentes restantes a regiones inferiores muy acotadas entre 0 y 40 aproximadamente. El resultado de las reglas 1,2 y 3 son tonos puros del rojo, el verde y el azul.

Las reglas 4 y 5 dan oscuridad y brillo a la imagen respectivamente. Cuando se detecta color en un pixel donde sus tres componentes poseen valores muy bajos, correspondientes a las funciones de pertenencia RojoOpaco, VerdeOpaco y AzulOpaco, se llevan a rangos menores expresados en las funciones de pertenencia RojoOpacoPsi, VerdeOpacoPsi



Tabla I Reglas de inferencia para el sistema difuso								
N.	Antecedente					Consecuente		
	Rojo	Op.	Verde	Op.	Azul	Rojo	Verde	Azul
1	Encendido	and	Opaco	and	Opaco	Encendido	Opaco	Opaco
2	Opaco	and	Encendido	and	Opaco	Opaco	Encendido	Opaco
3	Opaco	and	Opaco	and	Encendido	Opaco	Opaco	Encendido
4	Opaco	and	Opaco	and	Opaco	Opaco	Opaco	Opaco
5	Encendido	and	Encendido	and	Encendido	Encendido	Encendido	Encendido
6	Medio	and	Medio	and	Encendido	Opaco	Opaco	Encendido
7	Opaco	and	Encendido	and	Encendido	none	Medio	Encendido
8	Medio	and	Opaco	and	Opaco	Encendido	Medio	Opaco
9	Medio	and	Opaco	and	Opaco	Encendido	Opaco	Opaco
10	Encendido	and	Medio	and	Medio	Encendido	Opaco	Medio
11	Medio	and	Medio	and	Opaco	Encendido	Encendido	Opaco
12	Medio	and	Opaco	and	Encendido	Encendido	Opaco	Encendido
13	Medio	and	Encendido	and	Opaco	Opaco	Encendido	Opaco
14	Encendido	and	Encendido	and	Opaco	Encendido	Encendido	Opaco
15	Encendido	and	Opaco	and	Encendido	Encendido	Opaco	Encendido
16	Medio	and	Opaco	and	Medio	Encendido	Opaco	Encendido
17	Opaco	and	Medio	and	Medio	Opaco	Encendido	Encendido

y AzulOpacoPsi, lo que genera colores cercanos al negro permitiendo resaltar los colores generados por el resto de reglas. En el caso contrario, donde todas las componentes poseen valores altos correspondiente a colores claros se transforman en colores muy cercanos al blanco mediante los rangos determinados para las funciones de pertenencia de las variables de salida RojoEncendidoPsi, VerdeEncendidoPsi y AzulEncendidoPsi.

4. Ajuste de los conjuntos de pertenencia

Dado que se usó un criterio visual para determinar los rangos de las funciones de pertenencia para cada variable del sistema, los datos obtenidos son ambiguos pero cercanos a las características de color del estilo artístico. Por lo anterior, se contempló optimizar el sistema difuso propuesto mediante el método del gradiente.

La optimización del sistema busca generar valores para la media de la distribución y la desviación típica que describen las funciones de pertenencia para cada variable de entrada y salida, que son las variables que definen la función a optimizar, manteniendo el mismo número de reglas. Este proceso utiliza como información de entrada, una función objetivo que establece el error entre: los datos reales y los datos simulados. Los datos reales se obtuvieron mediante el concepto de expertos en el campo del arte¹, quienes en consenso aplicaron casos particulares de modificación de color con base en el arte psicodélico, permitiendo cuantificar los valores de referencia para una posterior comparación. Con la distinción de 72 colores expresados en código RGB (considerados como una muestra significativa que abarcan tonalidades oscuras, medias y encendidas), ubicados a lo largo de todo el espectro de color y que ejemplifiquen la variedad de los posibles valores de

1 Expertos consultados: el profesor de planta de la Facultad de Arte de la Universidad Distrital, Ricardo Forero Cárdenas y el profesor de honores de la Facultad de Arte de la Universidad Distrital, Alberto Díaz Lara.

entrada del sistema, los expertos generaron su equivalente en el estilo artístico. Los datos simulados se tomaron de las salidas del sistema de inferencia difusa descrito previamente cuya entrada fueron los casos de prueba mencionados.

En el Figura 4 se muestra el esquema general para realizar un ajuste sobre el sistema de inferencia difusa. Inicialmente los datos simulados ingresan como respuesta del sistema a la función objetivo, para calcular el error en comparación con los que se presenta entre los datos reales. Seguidamente el método de optimización busca disminuir el error generando, valores alternativos para los parámetros (media y varianza típica), tomándolos posteriormente con una nueva entrada de la función objetivo, para luego evaluarlos en el sistema difuso y obtener así un error ajustado, el cual puede ser sujeto a una nueva iteración de minimización.

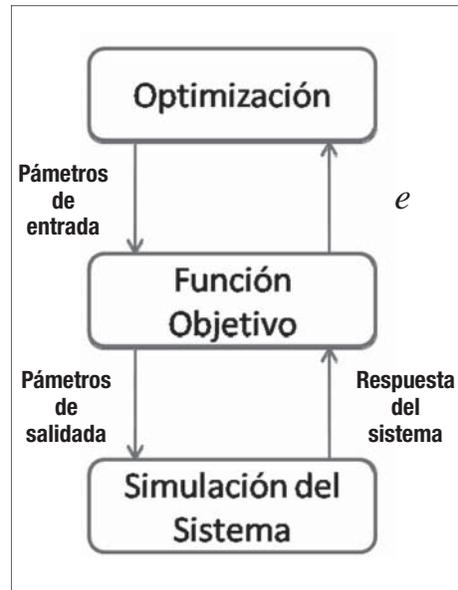


Figura 4. Esquema del proceso de optimización.

La función objetivo utilizada es el error cuadrático medio, debido a que la diferencia al cuadrado entre los datos reales y los datos simulados, permiten estimar errores tanto por defecto como por exceso. Esta función se define a continuación:

$$e_c = \sum_{i=1}^t (r_i - s_i)^2 \quad (1)$$

La variable e_c observada en la Ecuación (1) almacena el error (diferencia) entre los valores reales y simulados para el canal de color C , donde C corresponde a uno de los componentes del modelo de color RGB (rojo, verde o azul). Por su parte el índice i incrementa desde 1 hasta t que es el número de píxeles de la imagen digital a procesar. Las variables r_i y s_i corresponden a los valores reales y simulados (*respectivamente*) para el canal C ; el criterio del error cuadrático se aplica a cada uno de los componentes RGB por separado, de forma que, el error total se calcula sumando los errores obtenidos en cada canal, como se observa en la Ecuación (2).

$$e = e_R + e_G + e_B \quad (2)$$

Luego de definir la función objetivo, se decidió utilizar el método del gradiente (ilustrado en la Figura 5) para la optimización de los parámetros del sistema difuso. El método

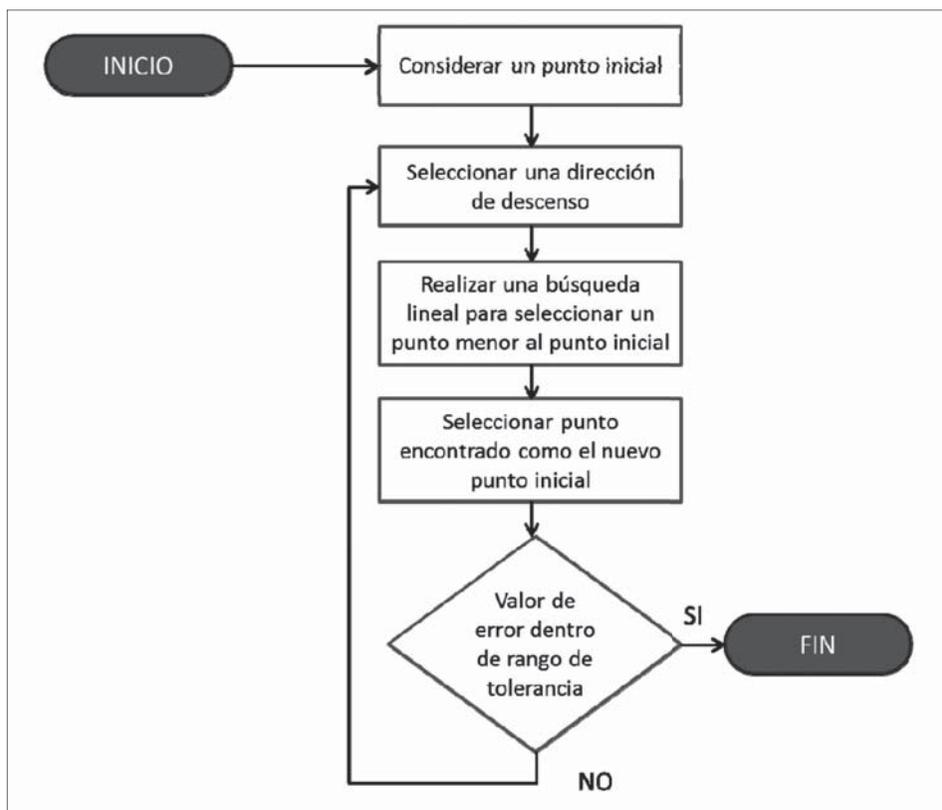


Figura 5. Método Gradiente

de optimización por el gradiente se aplica sobre funciones que se suponen continuas (aunque sea en intervalos). Estas funciones pueden contar con un único mínimo, con muchos mínimos locales, o con muchos mínimos locales y uno global. Como cualquier método de optimización para la minimización, el objetivo es encontrar el mínimo global, o en su defecto un mínimo local que cumpla con ciertos criterios aceptados en rangos de tolerancia.

El método del gradiente parte de un punto inicial a partir del cual se desciende en la dirección negativa del gradiente, esto es, el método genera puntos sucesivos mediante correcciones indicadas por el gradiente sobre los puntos generados previamente. El éxito de la aplicación del método en el proceso de encontrar un valor mínimo para la función proporcionada, depende en gran medida de la elección de un adecuado punto inicial; para el caso particular del sistema difuso propuesto, y teniendo en mente que los valores a optimizarse son los parámetros designados a cada función de pertenencia gaussiana (tanto la media como la desviación), el punto inicial se seleccionó con los valores determinados mediante el análisis visual dado para definición del sistema de lógica difusa previamente descrito.

```
x0= [51.39 25.5 14.8 165.2 31.25 257.2 54.9 16.29 15.61 149.1 45.53 246.9
     65.7 18.1 19.05 161 45.46 254.5 11.62 -0.767 4.519 128.2 4.408 254.3
     10.11 0.6632 4.743 125.6 8.011 256 9.812 -0.422 4.309 125.7 8.604 254];

options = optimset('Display','iter');
x = fminunc(@FuncionObjetivo,x0,options);
```

Figura 6. Ajuste del Sistema Difuso (Script en MATLAB).

Una vez definido el punto inicial, que consiste en un vector de 36 posiciones (número total de variables asociadas a las 18 funciones de pertenencia que incluyen las entradas y salidas) se ejecuta el método de la Figura 4 con la ayuda del software MATLAB® utilizando el *Toolbox* de optimización, mediante la funcionalidad de búsqueda del valor mínimo de la función objetivo (Ecuación (1)). Para este fin se utilizó la función `fminunc()`. Esta recibe como parámetros el punto inicial y la función de error. Internamente la función computa numéricamente la expresión correspondiente al gradiente (lo cual es transparente a quién utiliza el método) y un valor para el coeficiente que determina el tamaño del paso que se toma en la dirección contraria al gradiente en cada iteración. El script de MATLAB usado para realzar la optimización, se muestra en la Figura 6, donde las opciones (variable `options`) corresponden a la estructura de parámetros que recibe como entrada la función de optimización. Cada parámetro influye notablemente en la solución que obtenga el proceso de minimización. Los parámetros opcionales definidos mediante la instrucción `Optimset` se establecieron para mostrar la traza de la cantidad de iteraciones, así como su comportamiento. Las Ecuaciones (1) y (2) se implementaron en MATLAB mediante la siguiente expresión: `e=sum(sum(sum((r-s).^2)))`, la cual representa la función `@FuncionObjetivo`; donde `r` y `s` son las matrices de 72×3 (3 columnas correspondientes a cada componente del código RGB y 72 colores de muestra). Finalmente el punto inicial (`x0`) es el vector que contiene los parámetros de las funciones de pertenencia del sistema de inferencia original. El orden en que están organizado los parámetros en el punto inicial corresponde a parejas (desviación típica, media de la distribución) por cada conjunto, comenzando con las funciones de pertenencia de entrada (RojoOpaco, RojoMedio, RojoEncendido, VerdeOpaco, VerdeMedio, VerdeEncendido, AzulOpaco, AzulMedio, AzulEncendido) y continuando con las funciones de salida (RojoOpacoPsi, RojoMedioPsi, RojoEncendidoPsi, VerdeOpacoPsi, VerdeMedioPsi, VerdeEncendidoPsi, AzulOpacoPsi, AzulMedioPsi, AzulEncendidoPsi).

Se ejecutó el método durante 69 iteraciones hasta que no se continuaron realizando cambios significativos en el punto mínimo encontrado. La optimización calcula nuevos valores de media de la distribución y la desviación típica de cada función de pertenencia de las variables de entrada y salida, es decir genera un vector de 36 posiciones. Los nuevos valores de los parámetros ajustados de esta manera, se asignan a los conjuntos difusos para construir el sistema difuso mejorado. En la Figura 7 se observa el ajuste de las funciones de pertenencia optimizadas.

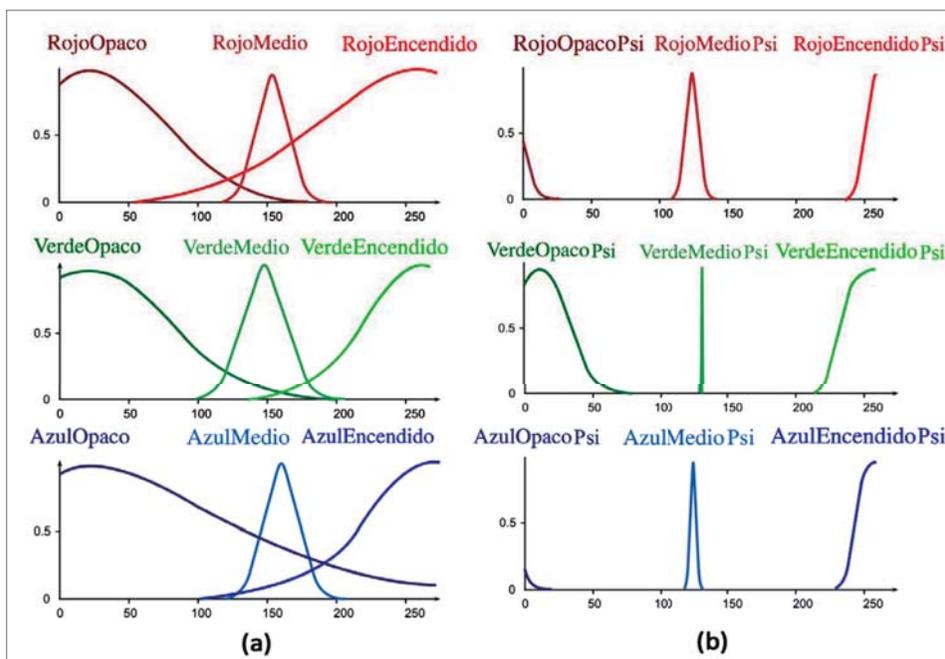


Figura 7. Funciones de pertenencia del sistema de inferencia optimizado.
(a) Variables de Entrada (color original) (b) Variables de salida (color psicodélico)

5. Resultados

Para realizar la coloración psicodélica se implementaron sistemas de inferencia difusa de tipo Mandani, utilizando el *Fuzzy Logic Toolbox* de MATLAB®. El primer sistema (sistema difuso original) se implementó utilizando las funciones de pertenencia mostradas en la Figura 3; el segundo (sistema difuso ajustado) utiliza las funciones de pertenencia de la Figura 5, optimizadas como se explica en la sección anterior. Ambos sistemas utilizan la base de reglas de inferencia de la Tabla I. Los experimentos se realizaron mediante un algoritmo de recorrido de cada pixel de la imagen de entrada, cuyos valores de color RGB son inyectados a los dos sistemas difusos, obteniendo finalmente la imagen de salida modificada al estilo psicodélico.

En la Tabla II se presenta el resultado de las salidas de los sistemas difusos para 11 imágenes digitales de cuadros con estilos artísticos diferentes al psicodélico. En la primera columna se observa la imagen original; en la segunda columna se observa la coloración utilizando el sistema difuso original; en la tercera columna se observa la coloración utilizando el sistema difuso ajustado.

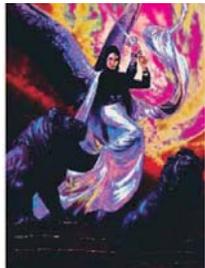
Una distinción inmediata de los resultados obtenidos con el sistema difuso ajustado, es que se destaca la continuidad del color y la conservación de los detalles de la imagen original, aspectos que no son evidentes en el sistema difuso original. Además se evidencia

un aumento en la intensidad del color y la alteración de colores como por ejemplo el paso de un gris rojizo a un morado saturado o de un naranja claro a un fucsia.

Una observación detenida de cada uno de los resultados presentados puede detallar diferentes casos donde varía la sensación visual de las imágenes resultantes. Para las imágenes originales 1, 2, 4, 5, 6, 8, 9, 10 y 11 (número de la columna izquierda de la Tabla II) se pueden percibir tonalidades de color muy tenues que son fuertemente resaltadas por los sistemas difusos (optimizado y sin optimizar), mientras muchas otras zonas, por la falta de brillo o claridad son llevadas a colores muy oscuros generando así contraste entre colores saturados y colores opacos.

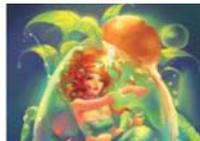
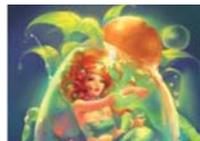
Las imágenes 1, 2, 4, 8, 9, 10 y 11 obtenidas con ayuda del sistema optimizado exponen una mayor cantidad de colores brillantes que resaltan de forma especial a la vista gracias al contraste con los fondos oscuros pudiéndose categorizar como más cercanas al estilo psicodélico, mientras en las imágenes 3, 5, 6 y 7 los efectos generados son más

Tabla II. Resultados del Sistema de Inferencia Difusa Original y Ajustado

Número	Imagen Original	Imagen Sin Optimizar	Imagen Optimizada	Créditos
1				Aralim - The Guardian Angel Artista: Volkan Baga © 2009
2				Tigre - Artista: Lourdes Marie Figueroa © 2009
3				Range op Eos - Artista: Volkan Baga © 2009



4				Teferi, Mage of Zhalfir - Artista: Volkan Baga © 2009
5				Wallpapers 3D Digital Art (93)
6				Wallpapers 3D Digital Art (8)
7				Peaceful place - Artista: Volkan Baga © 2009
8				Colección Seres de Fantasia. Sa-kimichan © blog oficial de la artista: http://sa-kimichan.tumblr.com/
9				Colección Seres de Fantasia. Sa-kimichan © blog oficial de la artista: http://sa-kimichan.tumblr.com/

10				Colección Seres de Fantasia. Sa-kimichan © blog oficial de la artista: http://sa-kimichan.tumblr.com/
11				Colección Seres de Fantasia. Sa-kimichan © blog oficial de la artista: http://sa-kimichan.tumblr.com/

sutiles, evidenciando así la necesidad de ampliar aún más el trabajo con nuevas reglas y estrategias de optimización.

En particular las imágenes 1 y 2 sin optimizar presentan sobresaturación del color, es decir colores directamente opuestos en el círculo cromático. La sobresaturación permite generar un efecto de realce o “movimiento” en la imagen, dadas las diferentes intensidades de color adyacentes. Los colores amarillo, fucsia y agua marina presentes en la imagen de forma contigua generan un mayor acercamiento al estilo artístico. Aun cuando las imágenes no presentan continuidad en el color y la permanencia en los detalles.

6. Conclusiones y expectativas futuras

La escala de colores del estilo artístico psicodélico se encuentra sobre tonalidades encendidas o tonos altos, dónde la transformación a dichas tonalidades se realiza mediante la saturación de colores que posean una o dos componentes en el código RGB con mayor valor numérico que las demás. Se observó que los dos sistemas de inferencia difusa construidos lograron capturar dichas características, aunque en mayor medida la aproximación al estilo psicodélico se logró con el sistema difuso ajustado, puesto que en este cambio de tonalidades se da de manera más continua, menos interrumpida, en términos artísticos, más natural. Lo anterior se logró como punto inicial de búsqueda del método de optimización, las funciones de pertenencia inicialmente propuestas de manera intuitiva en el sistema difuso original.

Un estudio interesante que podría desprenderse, sería realizar el ajuste utilizando puntos iniciales aleatorios para el método del gradiente, y comparar las variaciones obtenidas. Más aún se podría pensar en utilizar otros métodos de optimización, por ejemplo metaheurísticas basada en poblaciones de candidatos, como los Algoritmos Genéticos. En este último caso se tendría que considerar aspectos adicionales en el diseño de la



función objetivo (función de aptitud) y en la eficiencia del costo computacional requerido para su convergencia. Otra consideración de interés sería permitir que el Algoritmo Genético pueda ser alimentado con candidatos propuestos por diferentes artistas psicodélicos, de manera que el sistema se nutra de estilos personalizados.

Es importante resaltar que en este estudio solo se consideró la optimización de los parámetros de las funciones de pertenencia, manteniendo las mismas reglas de inferencia. En este aspecto surge otra pregunta en relación a si este conjunto de reglas también pueda ser susceptible a optimización, quizás para eliminar reglas irrelevantes, o para encontrar reglas más específicas, relacionadas con el estilo artístico psicodélico.

En cuanto a las tecnicidades del método utilizado, a futuro se busca modificarlo para la detección de vecindades contiguas, (no únicamente la lectura pixel a pixel) para sobresaturar la imagen digital. La sobresaturación consiste en la obtención de tonos puros o intensos por contraste de complemento, es decir, en el círculo cromático se define una relación entre tonos que se encuentran diametralmente opuestos, a 180 grados un tono del otro. La transformación que realiza el sistema difuso ajustado permitiría identificar el color en el estilo artístico psicodélico, para cambiar sus colores vecinos por colores opuestos. Esto permitiría que diferentes zonas de la imagen se vean resaltadas en intensidades de colores diferentes. En una imagen a trazos donde las regiones de un único color son más amplias, no se decidiría en la transformación de color sino en la asignación de colores que sobresaturen la imagen. Lo anterior tendría incluso aplicación en imágenes digitales en escala de grises en donde mediante la detección de zonas orientadas por el cambio en tonalidades, se generarían nuevas reglas para la selección adecuada el color que las transforme al estilo de arte psicodélico.

Otra vía de investigación alternativa puede ser expandir el campo de aplicación no solo a la generación de colores psicodélicos, sino a una primera aproximación sobre arte evolutivo, en cuanto a la modificación de la forma de las imágenes digitales. Estilos de arte visual como *op art* (optical art) y arte cinético podrían involucrarse con la saturación y sobresaturación del color. Este tipo de procesamiento podría en últimas incorporarse a las suites de software para tratamiento de imágenes digitales, de forma que permita cambiar el estilo artístico de la imagen en su totalidad o de secciones elegidas por el usuario, generando nuevos efectos visuales. Lo anterior podría ser de utilidad en el campo de la publicidad, diseño gráfico y afines. En este mismo sentido, la metodología seguida en este estudio, así como las consideraciones aquí discutidas, podrían eventualmente orientarse hacia otros estilos artísticos.

En cuanto a otras posibles aplicaciones, destacamos en primer lugar la coloración no solo en imágenes estáticas sino también en video y animaciones, constituyendo una nueva clase de filtro para la generación de efectos de un estilo artístico. Esto puede ser de utilidad en el campo del entretenimiento como el cine y televisión. Otras áreas que podrían beneficiarse incluyen el procesamiento de imágenes satelitales donde se podrían colorear psicodélicamente aquellos accidentes topográficos de interés, y en el campo médico, el

sistema propuesto podría generar casos de prueba de exámenes realizados a pacientes con problemas visuales de percepción de color como daltonismo. Se podría ir más allá y buscar que el sistema de manera dinámica adapte sus parámetros para generar transformaciones visuales adecuadas a la fisiología del paciente y a su representación interna del mundo y de esta manera diseñar tratamientos más efectivos.

7. Agradecimientos

Los autores desean agradecer a los artistas: Larry Carlson cuyo trabajo incluye una gran variedad de formas incluyendo fotografía, video-art, web-art, pinturas collage, arte digital, realización de películas, animación, text-art, y sound art; y de igual forma a la profesora de la Escuela y Galería de Artes Cruz, Inc. en Puerto Rico, Lourdes Marie Figueroa, por permitir el uso de su obra artística, pues su gentileza y oportuno interés apoyaron en gran medida el desarrollo del presente artículo.

Además, se agradece a al profesor de planta de la Facultad de Arte de la Universidad Distrital, Ricardo Forero Cárdenas y al profesor de honores de la Facultad de Arte de la Universidad Distrital, Alberto Díaz Lara, por su orientación en el campo artístico y manejo de color digital.

8. Referencias

- [1] Gabora, L. y DiPaola, S. (2007). Incorporating Characteristics of Human Creativity into an Evolutionary Art Algorithm. 10(2), 97-110
- [2] Yang Li. (2012). Adaptive Learning Evaluation Model For Evolutionary Art. WCCI 2012 IEEE World Congress on Computational Intelligence. 1 – 8.
- [3] Davoudi, M. y SeifNaraghi, N. (2009) Adaptive Subtitle and Caption Coloring Using Fuzzy Analysis. World Congress on Computer Science and Information Engineering. 764 - 768.
- [4] Celik, T., Özkaramanl, H., Demirel, H. (2007). Fire Pixel Classification Using Fuzzy Logic And Statistical Color Model. IEEE International Conference on Acoustics, Speech and Signal Processing. 1(1). I-1205 - I-1208
- [5] Millet, c., Bloch, I., Popescu, A. (2007). Using the Knowledge of Object Colors to Segment Images and Improve Web Image Search. Le Centre De Hautes Etudes Internationales D'informatique Documentaire. ACM Special Interest Group on Information Retrieval. 569-584.
- [6] Yuksel, M. y Basturk, A. (2012). Application of Type-2 Fuzzy Logic Filtering to Reduce Noise in Color Images. IEEE Computational Intelligence Magazine. 7(3) 25 – 35.
- [7] Chung, F. y Fung, B. (2003). Fuzzy Color Quantization and its Application to Scene Change Detection. Conference International Multimedia Conference ACM. New York, NY, USA.
- [8] Borji, A. y Hamidi, M. (2007). Evolving a Fuzzy Rule-Base for Image Segmentation. International Journal of Intelligent Technology. 2 (3) 471 – 476.
- [9] Khanale, P. y Kurhe, A (2011). Color Perception of Images Using Fuzzy Logic. Advances in Computational Sciences and Technology. 4(1) 1-8.



- [10] Jung-Min, O., Bang, B., Lee, G. (2008). Personal Color Decision System Using Fuzzy Logic. International Conference on Convergence and Hybrid Information Technology. Daejeon. 790 – 795.
- [11] Universidad de Extremadura. (2012). Aplicaciones de la lógica difusa a la colorimetría. Revista Hiperenciclopédica de Divulgación del Saber, 6(4).
- [12] Terano, T., Masui, S., Watanabe, H. (1992). Coloring of a Landscape by Fuzzy Logic. IEEE International Conference on Fuzzy Systems, 13 – 20.
- [13] Melomics y las nueva inteligencia computacional, canal universitario la opniondemalaga.es. (2012). [en línea]. Universidad de Malaga. Disponible en: <http://www.opinionmalaga.com/especiales/uma/2012/08/08/melomics-y-las-nueva-inteligencia-computacional.html> [2013, 01 de Junio]
- [14] Luft, Eric. (1952). Die at the Right Time! A Subjective Cultural History of the American Sixties. United States: United Book Press.
- [15] D’Negri, C. y De Vitro, E. (2006). Introducción al Razonamiento Aproximado: Lógica Difusa. Revista Argentina de Medicina Respiratoria. 4, 126 – 136.

Linda Sheriyn Rodríguez Castro

Estudiante Ingeniería de Sistemas de la Universidad Distrital Francisco José de Caldas, de Bogotá, Colombia. e-mail: lsrodriguez@correo.udistrital.edu.co

Iván Javier Ponce Fajardo

Estudiante Ingeniería de Sistemas de la Universidad Distrital Francisco José de Caldas, de Bogotá, Colombia. e-mail: ijponcef@correo.udistrital.edu.co

Helbert Eduardo Espitia Cuchango

Ingeniero Electrónico, Universidad Distrital Francisco José de Caldas. Ingeniero Mecatrónico, Universidad Nacional de Colombia. Especialista en Telecomunicaciones Móviles, Universidad Distrital Francisco José de Caldas. Magister en Ingeniería Mecánica, Universidad Nacional de Colombia. Docente de Planta, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. e-mail: heespitiac@udistrital.edu.co

Extensión de taxonomía y tratamiento de valores faltantes sobre un repositorio de incidentes de seguridad informática

Taxonomy extension and missing-values treatment over an informatics-security incident repository

Carlos Javier Carvajal Montealegre

Ingeniero de Sistemas
Universidad Distrital
Francisco José de Caldas
Ing.carlosj@gmail.com

Diego Nicolás Bayona

Ingeniero de Sistemas
Universidad Distrital
Francisco José de Caldas
Nicolas.bayona@gmail.com

Zulima Ortiz Bayona

Msc. en Matemáticas
Grupo de Investigación
Arquisoft, rama de
Seguridad Informática,
Universidad Distrital
Francisco José de Caldas
zorliz@udistrital.edu.co



Fecha recibida: ene. 18/2013
Fecha modificado: may. 27/2013
Fecha aceptado: may. 31/2013

Resumen

En este artículo se detalla el proceso de estimación de datos faltantes mediante el teorema de Bayes, sobre un repositorio de incidentes de seguridad informática compuesto por datos de tipo categórico. Así mismo, se hace uso de una taxonomía, ampliada y redefinida para acoplarse a los incidentes encontrados.

Palabras clave: datos categóricos, incidente de seguridad, minería de datos, pre-procesamiento de datos, repositorio, taxonomía.

Abstract

This paper describes the missing-values estimation process through the Bayes theorem acting over an information security incident repository composed by categorical data. Additionally, an augmented taxonomy is defined to account for the identified incidents.

Key words: categorical data, data mining, data pre-processing, security incident, repository, taxonomy.

1. Introducción

Un incidente de seguridad se define como la agrupación de una o más acciones llevadas a cabo por un atacante para lograr un resultado no autorizado en un sistema informático. Pueden ser distinguidas de otro grupo de acciones por las características de quien ataca, como ataca, qué objetivos tiene, quien fue el blanco del ataque y cuando se ejecutaron dichas acciones.



Este tipo de incidentes son reportados y dichos reportes son comúnmente accesibles desde Internet; en algunas ocasiones los reportes mantienen una estructura diferenciando algunas características previamente definidas, en otros casos son descripciones vagas y subjetivas. El conjunto de incidentes debe asegurar que sus elementos puedan ser descritos mediante las mismas características, para lo cual es necesario contar con un estándar descriptivo o taxonomía. Lo anterior con el fin de permitir análisis de minería de datos.

La taxonomía propuesta por el instituto Sandia en el artículo “A Common Language for Computer Security Incidents” [1], permite identificar y caracterizar estos incidentes, los campos de la taxonomía se observan en la figura 1.

Durante la recopilación de incidentes de seguridad realizada como parte de la investigación en el proyecto de grado “Análisis de Incidentes de Seguridad Informática Mediante Minería de Datos, para Modelado de Comportamiento y Reconocimiento de Patrones” [2] se observó la necesidad de ajustar la taxonomía antes descrita. De la misma forma, fue

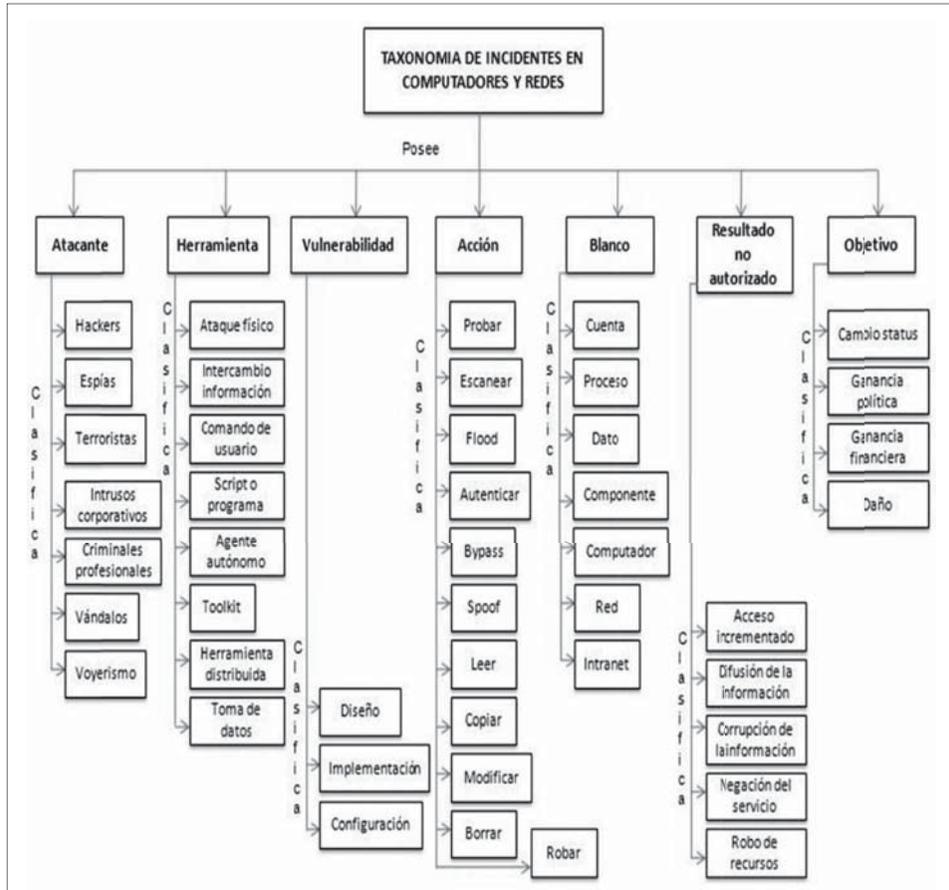


Figura 1. Taxonomía de incidentes de redes y computadores.

necesario considerar el problema de valores faltantes (missing values), en la construcción del repositorio, debido a que algunos valores de los incidentes no pueden ser capturados. Los resultados obtenidos en dicho proyecto se reportan en este artículo.

La organización del artículo es la siguiente: en la sección 1 se describe la obtención de los datos, para la Sección 2 se explica cómo y porque se extendió la taxonomía base, la Sección 3 muestra el procedimiento para completar los valores faltantes del repositorio, la sección 4 expone las pruebas y la sección 5 presenta los resultados obtenidos y las conclusiones.

2. Obtención de Datos

Se consultaron fuentes de reportes de incidentes, que fueran de acceso libre en Internet; es normal que las empresas que han sido blancos de ataques informáticos deseen mantener cierto nivel de reserva sobre las características del ataque, muchas veces mantienen en secreto la ocurrencia del incidente. Por otra parte los incidentes reportados en un sitio Web suelen ser los mismos citados en otros sitios, estos factores minimizan el número de ataques que pueden reunirse, no obstante se identificaron tres fuentes de fácil consulta y que proporcionaban gran número de datos sobre los incidentes:

Web Hacking Incident DataBase: Es un proyecto del Consorcio de Seguridad de Aplicaciones Web (Web Application Security Consortium) [3], dedicado a mantener una lista de incidentes de seguridad informática en aplicaciones Web. En esta se incluyen solo ataques dirigidos, aquellos que violaron la seguridad de las aplicaciones web. El sitio cuenta con filtros de búsqueda capaces de consultar los incidentes registrados, como: método de ataque, debilidad de la aplicación Web, resultado y entidad atacada. De esta fuente se adquirieron 569 incidentes con fecha de ocurrencia fuera mayor al 2007. Las características de los incidentes de esta base de datos se resumen en la Tabla I.

Chronology of Data Breaches: *Open Security Foundation*, es fuente de información acerca de violaciones de seguridad. La base de datos incluye datos como: el tipo de violación de seguridad, el tipo de entidad atacada, la ubicación geográfica de la entidad y la descripción del ataque [4]. Se tomaron 579 registros ocurridos entre el 2005 y el 2010, teniendo como condición que la violación de seguridad fuera HACK (Hacking o Malware) o DISC (Divulgación no intencional). Las características de los incidentes de esta base de datos se resumen en la Tabla II.

COL-CSIRT: Es un proyecto de investigación de la Universidad Distrital [5] con el fin de crear un equipo de respuesta a incidentes de seguridad informática. El grupo de investigación cuenta con una base de datos del centro de respuesta a incidentes y ataques con código malicioso; esta información proviene de las siguientes fuentes: Centro de estudios de respuesta y tratamiento de incidentes de seguridad de Brasil [6], y Centro de respuesta a incidentes de seguridad de Inteco [7]. Se obtuvieron 169 incidentes desde febrero de 2008 a octubre de 2010.

**Tabla I. WEB HACKING INCIDENT DATABASE**

Campo	Descripción
WHID-ID	Hace referencia al consecutivo registrado para cada incidente de seguridad informática almacenado en el repositorio. El formato definido para el campo es el siguiente WHID AAAA-Consecutivo.
Entry Title	Corresponde al Título con el que fue registrado el incidente de seguridad informática en la base de datos: WHID AAAA-Consecutivo: Sumario del incidente.
Incident Description	Corresponde a la descripción del incidente de seguridad informática, donde registra los detalles de la incidencia.
Reference	Enlaces a sitios web donde se encuentra almacenada información referente al incidente de seguridad informática en mención.
Date Occurred	Registra la fecha en la que ocurrió el incidente de seguridad informática. El formato de la fecha es el siguiente: MM-DD-AA.
Attack Method	Corresponde al método utilizado para realizar el ataque.
Application Weakness	Referencia a la debilidad vulnerada en la aplicación que permitió el éxito del ataque.
Outcome	Registra la finalidad del ataque realizado, es decir, lo que el ataque obtuvo como resultado para el atacante.
Attacked Entity Field	Registra el tipo de organización de la entidad atacada.
Attacked Entity Geography	Corresponde al sitio geográfico donde sucedió el incidente. El campo contiene el nombre de la ciudad atacada o en su defecto el nombre del país atacado.
Mass Attack	Indica si el ataque fue masivo. El campo registra la palabra Yes o No, informando si el ataque descrito en el Incidente de Seguridad Informática fue de tipo masivo.
Mass Attack Name	Registra el nombre del ataque masivo realizado.
Number of Site Affected	Número de sitios afectados en los ataques realizados en el incidente de seguridad. El formato del campo es un número entero que informa la cantidad de sitios atacados.
Attack Source Geography	Referencia al sitio geográfico de donde se realizó el ataque. El formato del campo es la dirección URL de la ubicación en el sitio web Google Maps.
Attacked System Technology	Indica el tipo de tecnología atacada.
Cost	Indica el costo monetario para la(s) entidad(es) atacada(s).
Items Leaked	Tipo de información robada en el ataque descrito en el Incidente de Seguridad informática.
Numbers of records	Indica el número de ataques realizados en el incidente registrado. El formato para el campo indica el número de ataques realizados en el incidente de seguridad informática. Para los casos que no se conoce el número de ataques, registra el valor "UNKNOWN".
Additional link	Almacena enlaces adicionales donde se registra información del incidente. El campo registra las direcciones URL adicionales que brindan información del incidente.

Tabla II. CHRONOLOGY OF DATA BREACHES	
Campo	Descripción
Date	Fecha de suceso del incidente de seguridad informática. El formato del campo corresponde al siguiente tipo de información "Mes Día de Año".
Name Entity	Nombre de la entidad atacada en el incidente.
Type Entity	Tipo de organización de la entidad víctima del incidente (Institución Educativa, Financiera, etc).
Type of Breach	Tipo de violación de seguridad el cual fue ejecutado en el incidente de seguridad informática (Hack, Disc).
Total Records	Cantidad de ataques realizados para el incidente en cuestión.
Description	Descripción de los detalles de la ocurrencia del incidente.
Geography Ubication	Ubicación geográfica donde el incidente de seguridad fue efectuado. El valor del campo indica la ciudad del suceso.

Los datos de estas tres fuentes se estandarizaron según la taxonomía Sandía mencionada anteriormente. Para ello se recorrieron de manera manual y se llenaron los campos definidos en la taxonomía.

3. Extensión de la Taxonomía

Luego de leer, analizar y clasificar los incidentes, se encontró dificultad en completar campos como la Herramienta, el Objetivo, el Atacante y el Resultado no autorizado, pues esta información se extrajo de la descripción proporcionada, que en más del 50% de los casos no era lo suficientemente explicativa. Para esos incidentes fue necesario una consulta adicional en Internet, en distintos artículos para tener mayor conocimiento sobre los ataques y completar los datos de clasificación [8][9][10][11][12][13].

Por otra parte cabe resaltar que los campos extras de la WHID almacenan información de gran interés para identificar los incidentes, entre los que se cuentan:

Debilidad de la Aplicación: Especifica el fallo que permitió la ocurrencia del incidente, con los siguientes posibles valores: Autenticación Insuficiente, Autorización insuficiente, Entropía insuficiente, Falta de Configuración del Antivirus, Falta de configuración de Infraestructura, Falta de configuración en el Servidor Web, Falta de configuración en la base de datos, Falta de configuración de la aplicación, Funcionalidad vulnerable, Indexado Inseguro, Insuficiente anti-automatización, Insuficiente expira-



ción de sesión, Insuficiente recuperación de contraseña, Manejo inapropiado de entradas, Manejo inapropiado de salidas, Permisos inapropiados sobre archivos del sistema, Protección insuficiente en la capa de transporte y Proceso de validación insuficiente.

Método de Ataque: Proporciona información detallada acerca de cómo se vulneró la seguridad y se obtuvo el acceso no autorizado, con los siguientes posibles valores: Abuso de funcionalidad, Automatización de procesos, Acceso no autorizado, ARP *Spoofing*, Bot, *Clickjacking*, Comando del sistema operativo, Cross Site Request Forgery (CSRF), Cross Site Scripting (XSS), Denegación de servicio, Desconocido, Divulgación no intencional, Fuerza bruta, Gusano, Hijacking DNS, Inclusión de archivo local, Inclusión de archivo remoto, Localización predecible de recursos, Navegación forzada, *Phishing*, Secuestro de dominio, Secuestro de sesión, *Spyware*, Software de rastreo, SQL *Injection*, Suplantación de contenido, Troyano y Virus.

Estos dos campos, por los valores que pueden tomar, dan un panorama más exacto del incidente y por tal motivo fueron incluidos en la taxonomía, generando una primer mejora.

A medida que se fueron analizando más incidentes se hicieron latentes nuevas necesidades, por una parte muchos de los valores de la taxonomía base (inclusive de los dos nuevos campos tomados de WHID) no abarcaban características esenciales de los incidentes y por otra parte, debido a la información que se suele reportar y encontrar sobre los incidentes, el campo Acción no podría ser diligenciado. Por las razones anteriores, en primer lugar se eliminó el campo Acción de la taxonomía y en segundo lugar se crearon nuevos valores para la clasificación, en las figuras 2 y 3 se muestra la taxonomía final extendida. En la Tabla III se resumen los nuevos valores creados para la taxonomía con su justificación. Estos a su vez, permitieron clasificar finalmente todos los incidentes encontrados cubriendo un rango más amplio que el de la clasificación existente, además de proporcionar mayor conocimiento del vector de ataque.

Tabla III. NUEVOS VALORES PARA TAXONOMÍA		
Valores	Campo	Justificación
Usuario Interno	Atacante	Incidentes causados por errores de empleados con acceso autorizado al sistema tal como el envío de correos a destinatarios no autorizados, negligencia al evitar el uso de contraseñas o instalación de software para fines de uso personal en computadores que albergan información privada de la entidad.
Políticas de Seguridad	Vulnerabilidad	
Sin Intencionalidad	Objetivo	
Usuario Externo	Atacante	Un usuario ajeno a la entidad, pero con un cierto nivel de acceso descubre por accidente errores o agujeros de seguridad que le permiten obtener privilegios más allá de los que tiene permitidos.

Bot	Método de Ataque	Incidentes donde el causante de la vulneración es un programa malicioso que busca infectar máquinas y convertirlas en miembros de redes zombie.
Spyware	Método de Ataque	Incidentes donde el software malicioso responsable de la vulneración está diseñado para recopilar anónimamente la información del equipo infectado y reenviar esta información a destinatarios específicos.
Software de Rastreo	Método de Ataque	Incidentes donde el ataque es realizado a través de software tipo escáner de Redes LAN; donde el programa analiza los puertos NetBios, las direcciones IP, las direcciones MAC, y puertos lógicos TCP/IP/UDP. El análisis es con el fin de encontrar vulnerabilidades que puedan ser aprovechadas por el atacante.
No Aplica	Herramienta	Indica que el incidente fue ocasionado por: un error, una falta de configuración, o la incorrecta validación de las vulnerabilidades de seguridad en los distintos componentes de Software o Hardware del sistema afectado.
No Aplica	Método de Ataque	Incidentes donde un usuario no tuvo que realizar ninguna acción especial fuera de las habituales para interactuar con el sistema, son ocasionados por que el sistema o componente es factible de atacar, debido a alguna vulnerabilidad en la configuración o en el incumplimiento de las políticas de seguridad informática.
No Aplica	Debilidad de la Aplicación	No existe debilidad en la aplicación, sujeto totalmente a incidentes donde existe un error humano. Publicaciones de información confidencial son el ejemplo más claro.
Falta de Configuración del Antivirus	Debilidad de la Aplicación	Incidentes donde un software de antivirus desactualizado permitió que un programa malicioso pudiera vulnerar la seguridad.
Falta de Configuración de la Infraestructura	Debilidad de la Aplicación	Se pasan por alto ciertas directivas sobre configuraciones de red, como evitar que las peticiones de internet lleguen a ciertas máquinas o tener un ancho de banda insuficiente para un servicio de alto flujo de datos.
Falta de configuración en el servidor WEB	Debilidad de la Aplicación	Por negligencia, los servidores Web se dejan con una configuración por omisión para contraseñas, estructura de directorios y nombres de carpetas entre otros, permitiéndole al atacante conocer exactamente donde están los archivos que desee acceder y como acceder a estos.
Falta de Configuración en Base de Datos	Debilidad de la Aplicación	Aunque las aplicaciones que accedan a información almacenada en bases de datos estén correctamente configuradas y sigan las directivas de seguridad, la base de datos carece de una contraseña por lo que la información puede ser accedida por otros medios diferentes a las aplicaciones determinadas para este fin.
Funcionalidad Vulnerable	Debilidad de la Aplicación	Un usuario descubre como evadir los controles de una aplicación mediante medios permitidos por la misma aplicación. [17].

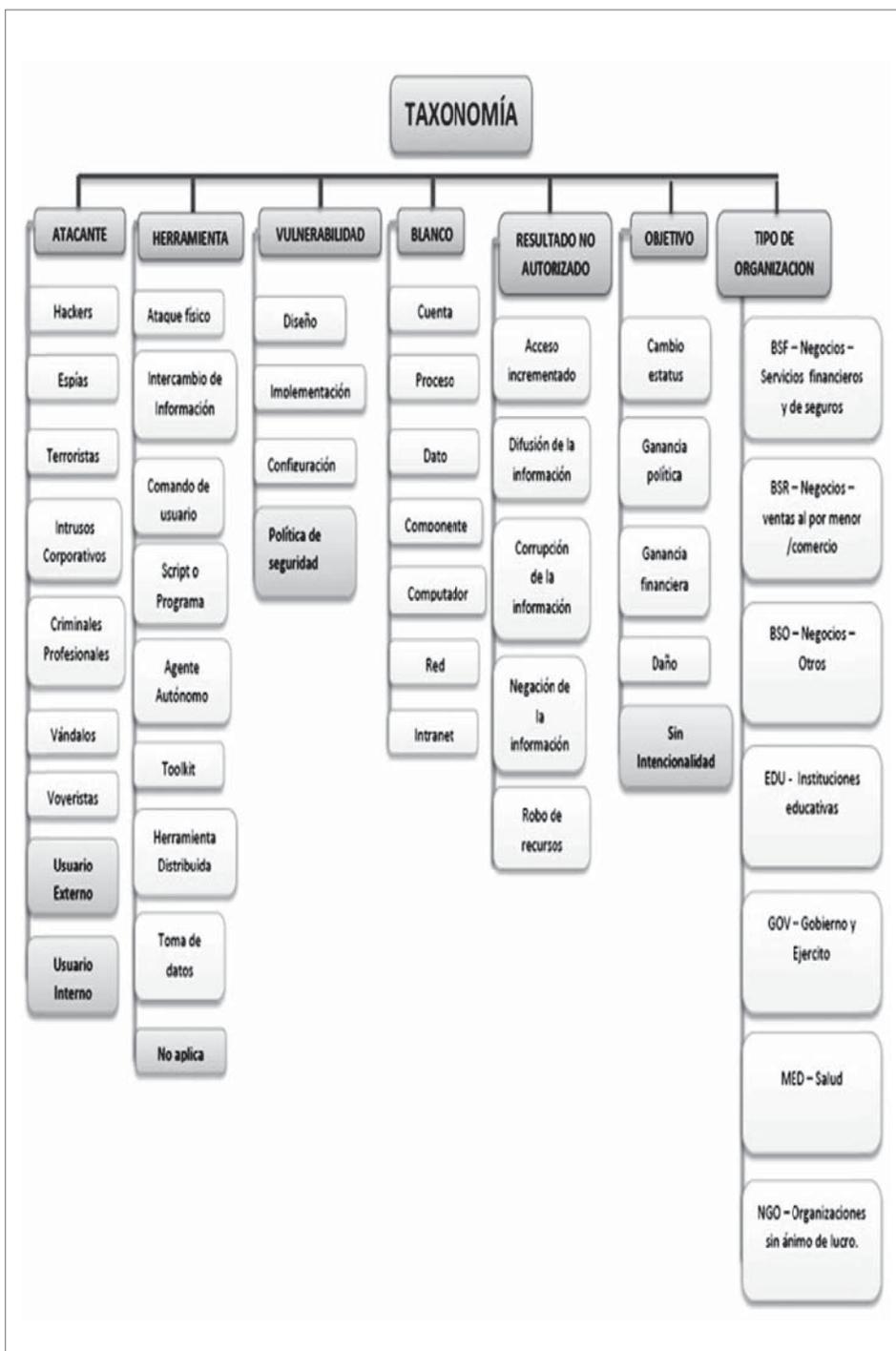


Figura 2. Taxonomía extendida de incidentes seguridad informática. Parte 1

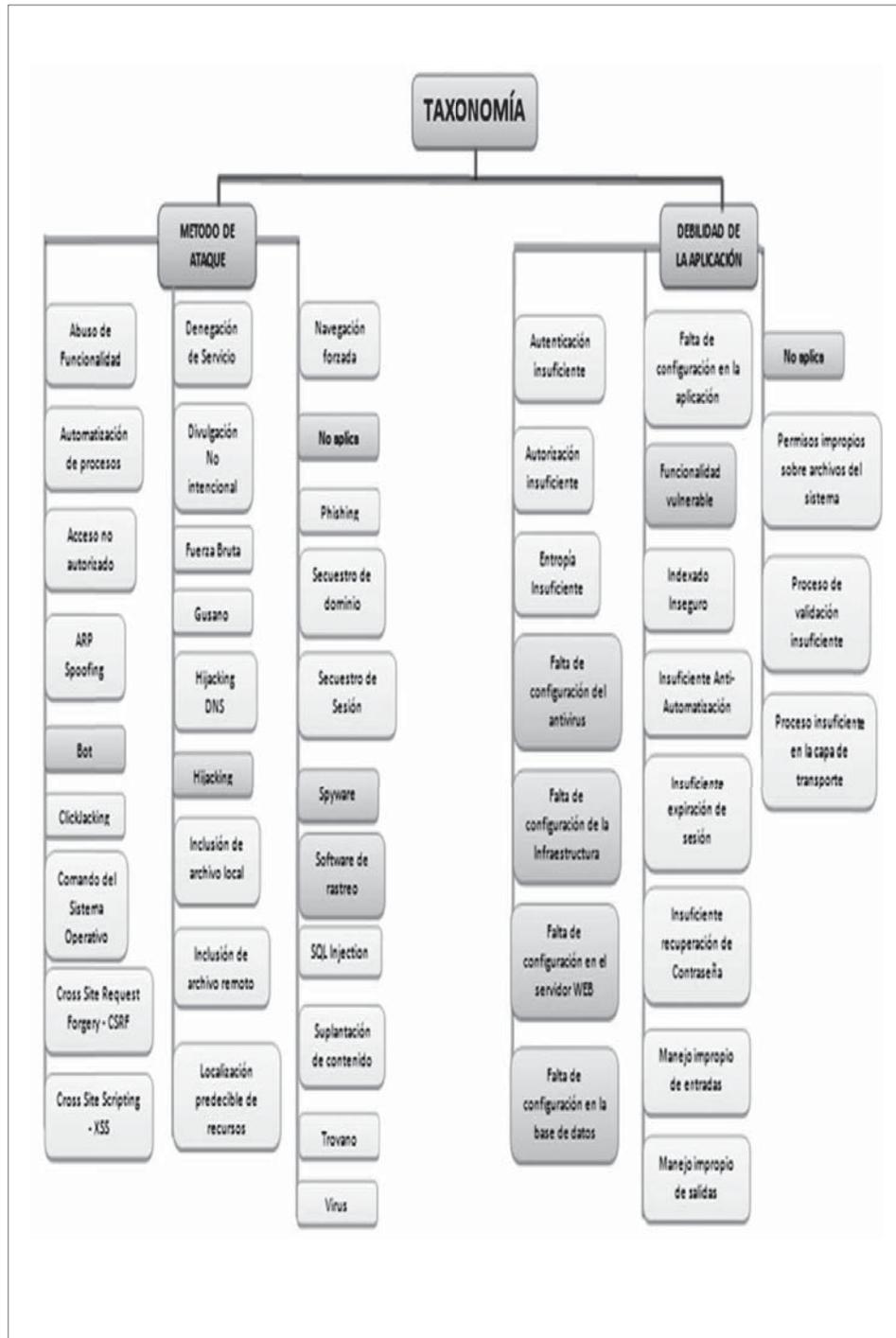


Figura 3. Taxonomía extendida de incidentes seguridad informática. Parte 2



4. Tratamiento de valores faltantes

Una vez finalizado el proceso de estandarización de datos, el número total de registros recopilados y descritos de acuerdo a la taxonomía mejorada fue de 1313 incidentes de seguridad, no obstante, debido a la precaria información relacionada para el 25% de los incidentes, no fue posible completar todos los campos descriptivos, este problema es más conocido en la literatura como *missing-values*, valores faltantes o en blanco. De los 1313 registros, 982 tuvieron todos los campos completos, mientras que 331 presentaron uno o más valores faltantes, los casos de mayor impacto se presentaron en los campos Método de Ataque 145 registros, Debilidad de la Aplicación 123 registros, Vulnerabilidad 144 registros y Herramienta 262 registros con valores faltantes. En las figuras 4 y 5 se puede observar la distribución de valores para los campos Método de Ataque y Debilidad de la Aplicación y se visualiza el alto número de registros con estos campos en blanco.

Tal como lo sugieren autores consultados [14][15][16], las opciones para tratamiento de *missing-values* son:

- Ignorar el registro cuando tenga gran cantidad de atributos faltantes.
- Llenar el valor faltante de manera manual.
- Usar una constante global para llenar el valor faltante, por ejemplo el texto “desconocido”.
- Usar el valor promedio que muestra el atributo en los registros.
- Usar el valor más probable determinado mediante herramientas basadas en inferencia, tales como el formalismo Bayesiano y árboles de decisión.
- Métodos de aprendizaje supervisado y no supervisado tales como el clustering, o las redes neuronales.
- Estas opciones se analizaron de manera individual:
- Ignorar los registros con valores faltantes conllevaría a perder el 25% de los incidentes recopilados.
- Llenar el valor faltante manualmente no es una opción pues el repositorio se construyó de esta manera.

Definir un valor por omisión de llenado causaría que estos algoritmos consideraran dicha constante como un valor de “interés”; calcular un valor promedio es una buena opción para repositorios cuyos campos sean numéricos, sin embargo en este caso son categóricos con lo cual no es posible hallar un promedio [17] [18]. Los métodos de aprendizaje supervisado y no supervisado serán usados a futuro en tareas propias de la minería de datos [19], tales como la detección de *outliers* y el análisis de cluster en el momento en que el repositorio este estandarizado totalmente. Dado este análisis, la estimación del valor más probable se escogió como la opción más conveniente para el caso de estudio. La técnica utilizada se define con detalle a continuación.

4.1 Estimación de valores faltantes mediante un enfoque Bayesiano.

Se utilizó una técnica con base en la propuesta de estimación de valores categóricos faltantes de Xiao-Bai Li [20], donde los valores faltantes de tipo categóricos se estiman como un problema de clasificación mediante cálculos probabilísticos, es decir se calcula el valor faltante del atributo de interés, como la clase (categoría) más probable dados los valores para los restantes atributos presentes. Esta probabilidad se obtiene mediante una aproximación Bayesiana:

Sea c_1, \dots, c_L particiones del espacio muestral. Entonces, para cada evento X en el espacio muestral,

$$P(c_k | X) = \frac{P(c_k)P(X | c_k)}{\sum_{r=1}^L P(c_r)P(X | c_r)}, k = 1, \dots, L, \quad (1)$$

Donde $P(c_k)$ es llamada la probabilidad a priori y $P(c_k | X)$ es la probabilidad a posteriori.

Se considera un conjunto de datos con un atributo de clase de dos tipos posibles, c_1 y c_2 , y $M-1$ atributos que no son de clase, X_1, \dots, X_{M-1} . Para un nuevo registro $x = (x_1, \dots, x_{M-1})$ que deba ser clasificado; el clasificador Bayesiano asigna al valor de clase a c_1 si $P(c_1 | x) > P(c_2 | x)$, de lo contrario asigna a c_2 . La probabilidad a posteriori $P(c_k | x)$ puede ser derivada del Teorema de Bayes (1). El proceso involucra la estimación de $P(c_k)$ y $P(x | c_k)$ a partir de los datos. Mientras que $P(c_k)$ es fácil de obtener, evaluar a $P(x | c_k)$ es muy costoso en cuanto a recursos computacionales para conjuntos de datos con una alta dimensionalidad. Para sortear este inconveniente, se asume que los atributos son condicionalmente independientes entre sí. Bajo esta suposición, $P(x | c_k)$ puede ser fácilmente calculado por:

$$P(x | c_k) = \prod_{j=1}^{M-1} P(x_j | c_k) \quad (2)$$

Lo que se conoce como un clasificador ingenuo de Bayes (Naive Bayes classifier).

El método permite estimar las probabilidades de múltiples atributos faltantes con base en los atributos existentes. Dado un conjunto de datos con N registros y M atributos categóricos, X_1, \dots, X_M , sea L_i el número de categorías en X_i , y sea N_i el número de registros con X_i valores conocidos, y N_{ij} el número de registros donde X_i es igual a la k -ésima categoría c_{ij} . Adicionalmente, sea $N_{jr} | i_k$ el número de registros donde X_j es igual a la r -ésima categoría c_{jr} , dado $X_i = c_{ik}$ con $j \neq i$. El proceso para completar los valores faltantes a partir de este método se describe en la Tabla IV.

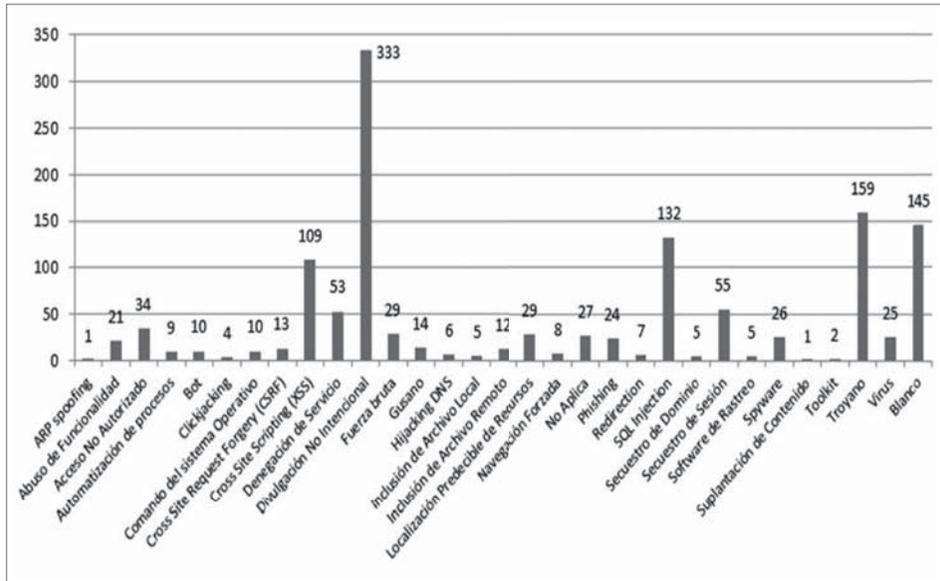


Figura 4. Distribución de valores Iniciales para Método de Ataque.

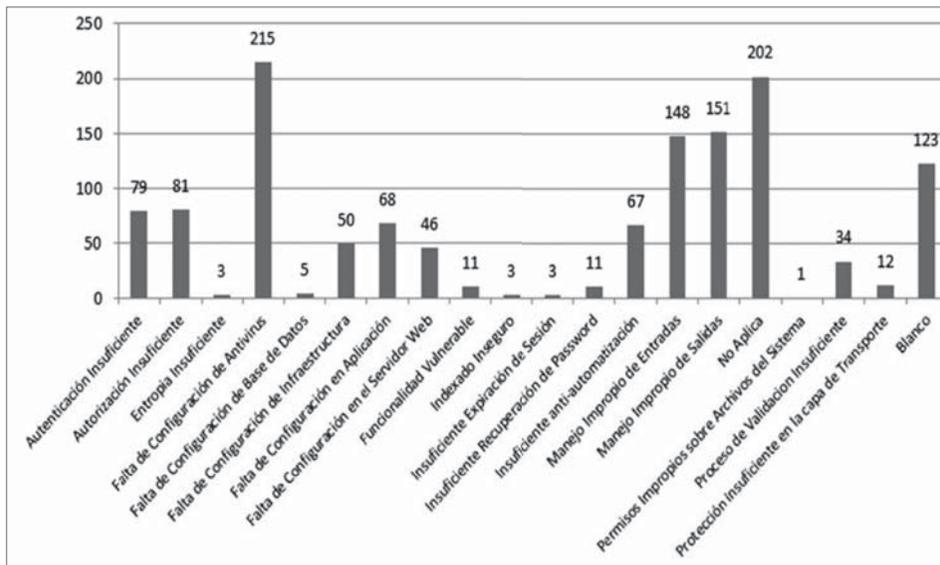


Figura 5. Distribución de valores Iniciales para Debilidad de la Aplicación.

4.2 Ejemplos de estimación

Consideremos una versión simplificado de nuestro repositorio, compuesto por 4 atributos categóricos: ATACANTE, HERRAMIENTA, VULNERABILIDAD y OBJETIVO.

Tabla IV. ESTIMACION DE VALORES FALTANTES	
Procedimiento	
1. Calcular las probabilidades a priori para cada atributo:	$P(X=C_{ik}) = N_{ik}/N_i, i = 1, \dots, m; k = 1, \dots, L_i.$
2. Calcular las probabilidades condicionales de X_j , dado $X_i=c_{ik}$:	$P(X_j=C_{jr} X_i=C_{ik}) = N_{jr ik} / N_{ik}, j = 1, \dots, M; j \neq i; r = 1, \dots, L_j.$
3. Para un registro x con un valor faltante en X_i , sea J el índice para el conjunto de atributos sin valores faltantes en x , x_j la parte correspondiente de x . Calcular las probabilidades a posteriori, según las ecuaciones (1) y (2):	$P(X_i = c_{ik} x_j) = \frac{1}{P(x_j)} P(X_i = c_{ik}) \prod_{j \in J} P(X_j = c_{jr} X_i = c_{ik}), k = 1, \dots, L_i$
4. Reemplazar el valor faltante X_i de x basándose en las probabilidades calculadas en el paso 3 usando el valor con la máxima probabilidad a posteriori. (MaxPost).	

Suponemos que el repositorio se compone de 20 registros, donde cuatro de estos registros presentan valores faltantes como se observa en la Tabla V, para los registros 17 y 18 se observa que el campo faltante es VULNERABILIDAD, en este caso iniciamos calculado las probabilidades a priori para el mismo:

$$P(\text{Diseño}) = 4/18$$

$$P(\text{Configuración}) = 7/18$$

$$P(\text{Políticas de Seguridad}) = 7/18$$

A continuación se calculan las probabilidades condicionales para el atributo ATACANTE cuando sea igual a Hackers dado cierto valor de VULNERABILIDAD:

$$P(\text{Hackers} | \text{Diseño}) = 4/4$$

$$P(\text{Hackers} | \text{Configuración}) = 4/7$$

$$P(\text{Hackers} | \text{Políticas de Seguridad}) = 2/7$$

De la misma forma calculamos las probabilidades condicionales para {Herramienta = Agente Autónomo} y {Objetivo = Daño}

$$P(\text{Agente Autónomo} | \text{Diseño}) = 1/4$$

$$P(\text{Agente Autónomo} | \text{Configuración}) = 6/7$$

$$P(\text{Agente Autónomo} | \text{Políticas de Seguridad}) = 3/7$$

$$P(\text{Daño} | \text{Diseño}) = 4/4$$

$$P(\text{Daño} | \text{Configuración}) = 3/7$$

$$P(\text{Daño} | \text{Políticas de Seguridad}) = 2/7$$

Luego se calculan las probabilidades a posteriori:



$$P(DI|H,AA,D) = 1/P (4/18)(4/4)(1/4)(4/4) = 0,05556/P$$

$$P(C|H,AA,D)=1/P (7/18)(4/7) (6/7) (3/7) = 0,08163/P$$

$$P(PS|H, AA,D) =1/P (7/18) (2/7) (3/7) (2/7) = 0,013605/P$$

DI=Diseño, C= Configuración, PS= Políticas de Seguridad, H=Hackers, AA = Agente Autónomo D = Daño

Finalmente calculamos:

$$P(DI|H,AA,D) = 0,05556P/0,05556P + 0,08163P + 0,013605P = 0,36844$$

$$P(C|H,AA,D) = 0,08163P/0,05556P + 0,08163P + 0,013605P = 0,54133$$

$$P(PS|H,AA,D) = 0,013605P /0,05556P + 0,08163P + 0,013605P = 0,09022$$

Así que el valor más probable y el que reemplazara al valor faltante en los registros 17 y 18 es Configuración, este mismo cálculo se podría llevar a cabo para reemplazar los faltantes de los registros 19 y 20.

El mismo proceso se llevó a cabo con el repositorio real para completar los valores faltantes.

Tabla V. Repositorio Simplificado para Ejemplo de Estimación

No	Atacante	Herramienta	Vulnerabilidad	Objetivo
1	Hackers	Agente Autónomo	Diseño	Daño
2	Criminales Profesionales	Agente Autónomo	Configuración	Ganancia Financiera
3	Hackers	Agente Autónomo	Políticas de Seguridad	Daño
4	Hackers	Agente Autónomo	Configuración	Daño
5	Hackers	Agente Autónomo	Configuración	Cambio de Status
6	Hackers	Agente Autónomo	Configuración	Daño
7	Criminales Profesionales	Agente Autónomo	Configuración	Ganancia Financiera
8	Criminales Profesionales	Agente Autónomo	Políticas de Seguridad	Ganancia Financiera
9	Hackers	Comando de Usuario	Diseño	Daño
10	Hackers	Agente Autónomo	Configuración	Daño
11	Hackers	Agente Autónomo	Políticas de Seguridad	Daño
12	Hackers	Comando de Usuario	Diseño	Daño
13	Hackers	Comando de Usuario	Diseño	Daño
14	Usuario Interno	Intercambio de Información	Políticas de Seguridad	Sin Intencionalidad
15	Usuario Interno	Intercambio de Información	Políticas de Seguridad	Sin Intencionalidad
16	Usuario Interno	Intercambio de Información	Políticas de Seguridad	Sin Intencionalidad
17	Hackers	Agente Autónomo	?	Daño
18	Hackers	Agente Autónomo	?	Daño
19	?	?	Configuración	Ganancia Financiera
20	Usuario Interno	?	Políticas de Seguridad	?

4.3 Exclusión de Atributos con Bajo Contenido de Información

Para la estimación Bayesiana realizada, la dimensionalidad del incidente fue reducida, es decir se retiraron las siguientes columnas:

- **ENTIDAD**, 993 valores distintos.
- **FECHA**, 842 valores distintos.
- **PAIS**, el 70% de los incidentes tiene como país de origen a Estados Unidos, el 11% a Colombia, el 1,83% al Reino Unido, 1,29% en India, 1,14% en Australia y 12,34% en otros países.

El motivo de esta reducción se debe a que las columnas **ENTIDAD** y **FECHA**, para fines prácticos son identificadores del registro, pues en los incidentes recopilados pocos son reincidentes para una misma **ENTIDAD** y pocos ocurrieron en una misma fecha, ya que estas columnas manejan un alto rango de valores, no son aptas para agrupar por subconjunto.

En el otro extremo tenemos el campo **PAÍS**, ya que la mayoría de los incidentes ocurrieron en Estados Unidos, por lo tanto si se considerara en análisis Bayesiano de los valores faltantes sesgaría la decisión hacia el subconjunto definido para ese país. Por las razones anteriores los cálculos de probabilidad de las columnas mencionadas no se tomaron en cuenta para la estimación Bayesiana aunque siguen haciendo parte del repositorio.

5. Validación del método

Con el fin de comprobar la efectividad en la estimación de valores faltantes mediante el enfoque Bayesiano y probar el grado de confianza en el repositorio; se realizó una validación contra los registros del repositorio que tenían los datos completos. La distribución final del repositorio se muestra en la Tabla VI:

Tabla VI. Repositorio de datos		
Tipo de Registro	Número de Registros	Porcentaje
Completo	982	74,79%
Con valores faltantes	331	25,21%

Tabla VII. Pruebas estimación valores faltantes		
	Cantidad	Porcentaje
Aciertos	203	69%
Errores	91	31%

De la porción del repositorio que no tenía valores faltantes, se eligió un subconjunto equivalente al 30% (294 registros). Para cada registro del subconjunto de validación, se eliminó una columna elegida al azar y los valores eliminados fueron calculados con el enfoque Bayesiano descrito, luego se compararon los resultados obtenidos con el valor que tenía el campo inicialmente. Los resultados globales se presentan en la Tabla VII, los resultados discriminados por campo en la Tabla VIII y Figura 6.



Tabla VIII. Pruebas por campo de estimación de valores faltantes				
Campo	Aciertos	Errores	Porcentaje Aciertos	Porcentaje de error
Tipo de organización	14	11	56,00%	44,00%
Método de ataque	28	15	65,12%	34,88%
Debilidad de la aplicación	23	11	67,65%	32,35%
Atacante	32	11	74,42%	25,58%
Herramienta	28	11	71,79%	28,21%
Vulnerabilidad	20	0	100%	0%
Blanco	20	9	68,97%	31,03%
Resultado no autorizado	21	13	61,76%	38,24%
Objetivo	21	6	77,78%	22,22%

Como se evidencia el porcentaje más bajo de aciertos corresponde al Tipo de Organización, demostrando que no existe una relación estrecha entre este campo y el resto del vector del ataque. Por el contrario la Vulnerabilidad depende totalmente del resto de campos del registro logrando un porcentaje de aciertos del 100%.

El clasificador Bayesiano muestra alta dependencia de la distribución de valores propia del conjunto de datos y obtiene valores acertados para registros que tengan un comportamiento común y conocido para el clasificador, esto se deduce de la inspección de

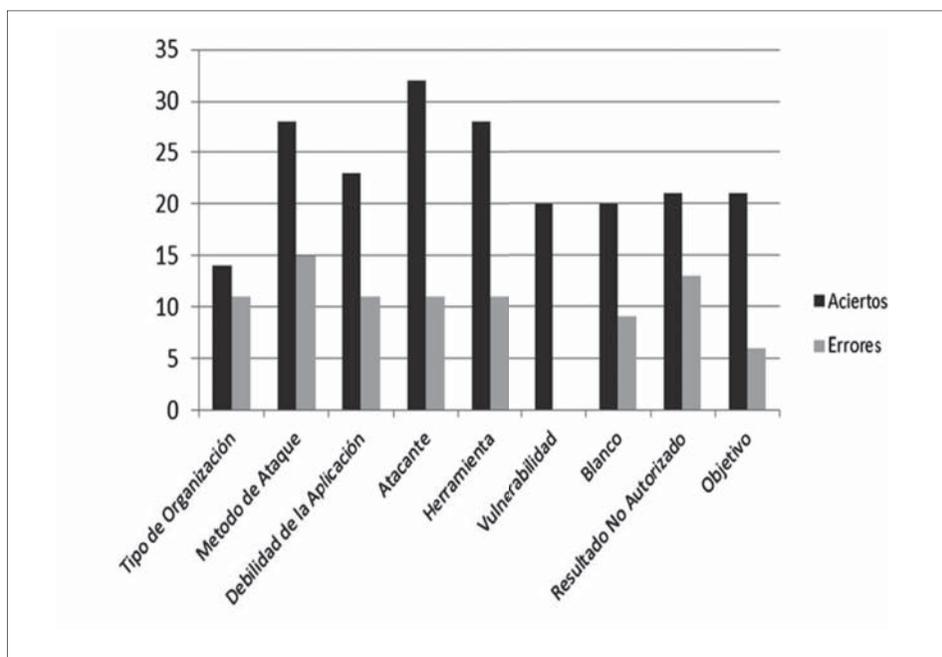


Figura 6. Cantidad de aciertos y errores mediante estimación Bayesiana

Tabla IX. Pruebas por campo de estimación de valores faltantes			
Campo eliminado	Valor esperado	Valor obtenido	Motivo del error
Método de Ataque	Secuestro de Sesión	No Aplica	Único incidente en el conjunto de datos de Secuestro de Sesión teniendo como blanco a Computador, ya que no existían más incidentes de este tipo la probabilidad tiende a cero.
Atacante	Usuario Interno	Hackers	Incidente donde un Usuario Interno uso un Troyano, siendo esto un comportamiento poco común.

algunos de los registros que no fueron clasificados correctamente como los que se presentan en la Tabla IX.

5.1 Comparación contra el valor promedio

El clasificador fue enfrentado con uno de los métodos más comunes en el tratamiento de los valores faltantes el cual consiste en elegir el valor promedio del campo para completar el registro. Los resultados totales del método del valor promedio se muestran en la tabla X y los resultados discriminados por campo se presentan en la Tabla XI y Figura 7.

Tabla X. Pruebas estimación valores faltantes		
	Cantidad	Porcentaje
Aciertos	101	34,35%
Errores	193	65,64%

Tabla XI. Pruebas por campo de estimación de valores faltantes				
Campo	Aciertos	Errores	Porcentaje Aciertos	Porcentaje de error
Tipo de organización	9	16	36,00%	64,00%
Método de ataque	8	35	18,60%	81,40%
Debilidad de la aplicación	6	28	17,65%	82,35%
Atacante	14	29	32,56%	67,44%
Herramienta	10	29	25,64%	74,36%
Vulnerabilidad	3	17	15,00%	85,00%
Blanco	20	9	68,97%	31,03%
Resultado no autorizado	18	16	52,94%	47,06%
Objetivo	13	14	48,15%	51,85%

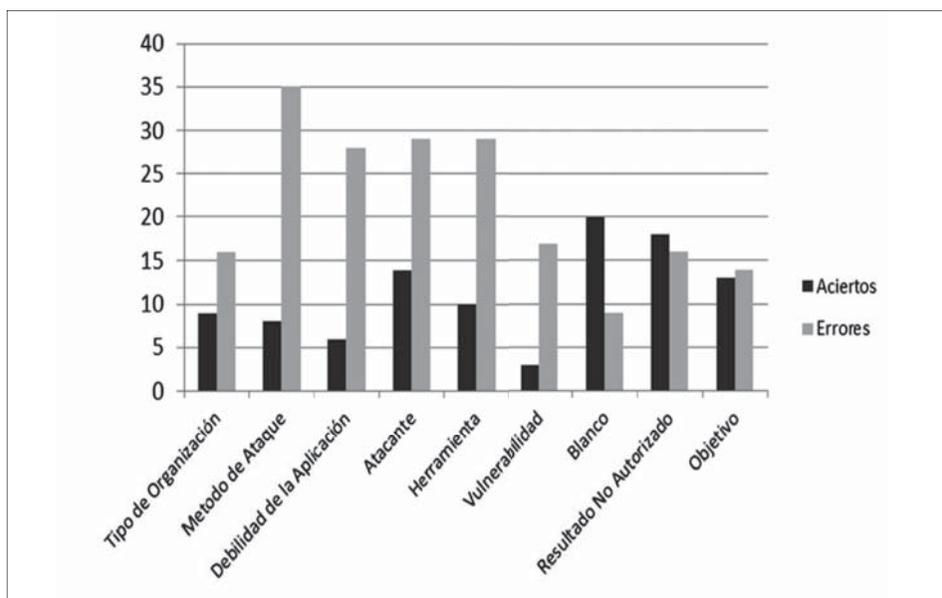


Figura 7. Cantidad de aciertos y errores mediante valor promedio

El método del valor por promedio muestra resultados menos efectivos en comparación a los del clasificador Bayesiano, el porcentaje de aciertos del clasificador Bayesiano (69%) es más del doble del porcentaje del método del promedio (34%) y además el valor promedio no logró sobrepasar en ninguno de los campos al competidor.

6. Resultados

Finalmente se obtuvo un repositorio de 1234 registros con todos los atributos diligenciados (Tipo de Organización, Entidad, Método de Ataque, Debilidad de la Aplicación, País, Fecha, Atacante, Herramienta, Vulnerabilidad, Blanco, Resultado No Autorizado y Objetivo), la técnica usada permitió completar el 76 % de los registros con valores faltantes, a pesar de la estimación Bayesiana, fue necesario eliminar 79 registros de los 1313 originales, pues éstos presentaban valores faltantes en las columnas **Método de Ataque**, **Debilidad de la Aplicación**, **Herramienta** y **Vulnerabilidad**, es decir una pérdida de información del 44% de la totalidad de atributos usados para hallar las probabilidades *a-posteriori*.

El repositorio final muestra (Figura 8) que los métodos de ataque predilectos son la Divulgación No Intencional (333 incidentes), los troyanos (161 incidentes), la inyección de SQL (134 incidentes) y el Cross Site Scripting (122 incidentes). El uso de Virus comunes se ve rezagado (34 incidentes), al igual que los gusanos (14 incidentes). Los métodos menos populares son Clickjacking y Toolkit (ambos con 4 incidentes), ARP spoofing y Suplantación de Contenido (1 incidente cada uno).

La Divulgación no Intencional fue causada en un 83% de los casos por Usuarios Internos, esto demuestra la insuficiente implantación de Políticas de Seguridad y el énfasis que debe aplicarse a la capacitación de los empleados sobre la importancia de la seguridad de la información. El alto uso de troyanos lleva a la misma conclusión, ya que en el 93% de los casos la debilidad de la aplicación fue Falta de Configuración en el Antivirus.

Para los incidentes de SQL Injection y Cross Site Scripting, la Vulnerabilidad fue Implementación en un 87% y 84% respectivamente, la responsabilidad de estos incidentes recae en los desarrolladores del software. La deficiencia en el desarrollo del software se centra en la codificación de la interfaz de usuario como se ve en la figura 9, con 152 incidentes causados por un Manejo Impropio de Salidas y 150 por Manejo Impropio de Entradas. El número de incidentes de este tipo debería ser tomado en cuenta en las empresas con el fin de analizar la capacidad de los desarrolladores contratados y tomar medidas de tipo gerencial, como capacitar o cambiar el equipo de desarrolladores.

En contra de una suposición popular, las entidades de tipo financiero obtuvieron el cuarto lugar de tipo de entidades atacadas, siendo las más vulneradas las que clasifican en la categoría BSO - Negocios - Otros, conformado por empresas de tecnología, medios de comunicación y hoteles entre otros. Por otra parte, las instituciones dedicadas a la salud son las menos atacadas. En la figura 10 se observa la distribución de incidentes por categoría.

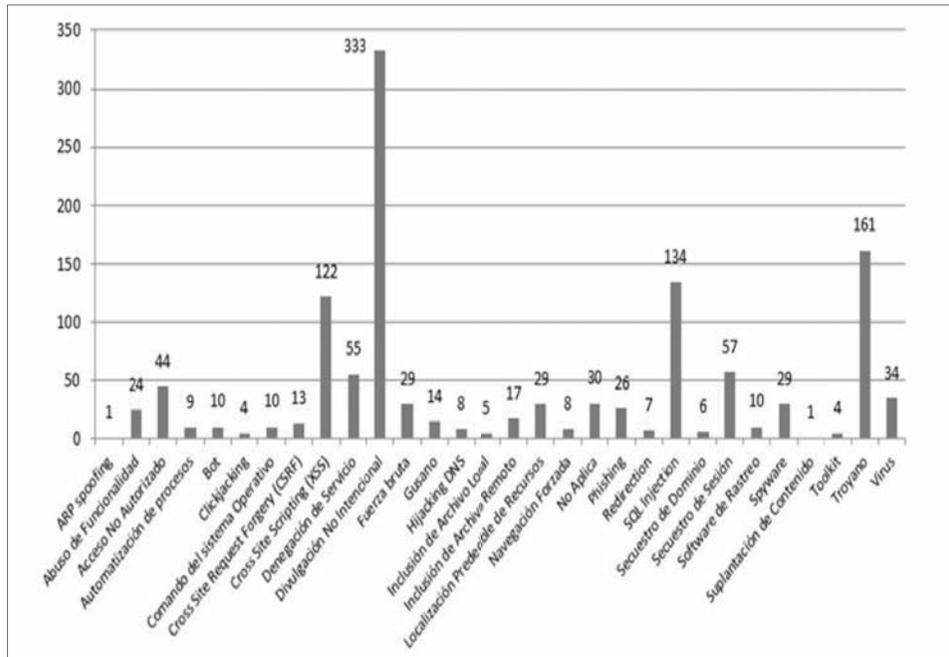


Figura 8. Distribución final de Valores Método de Ataque

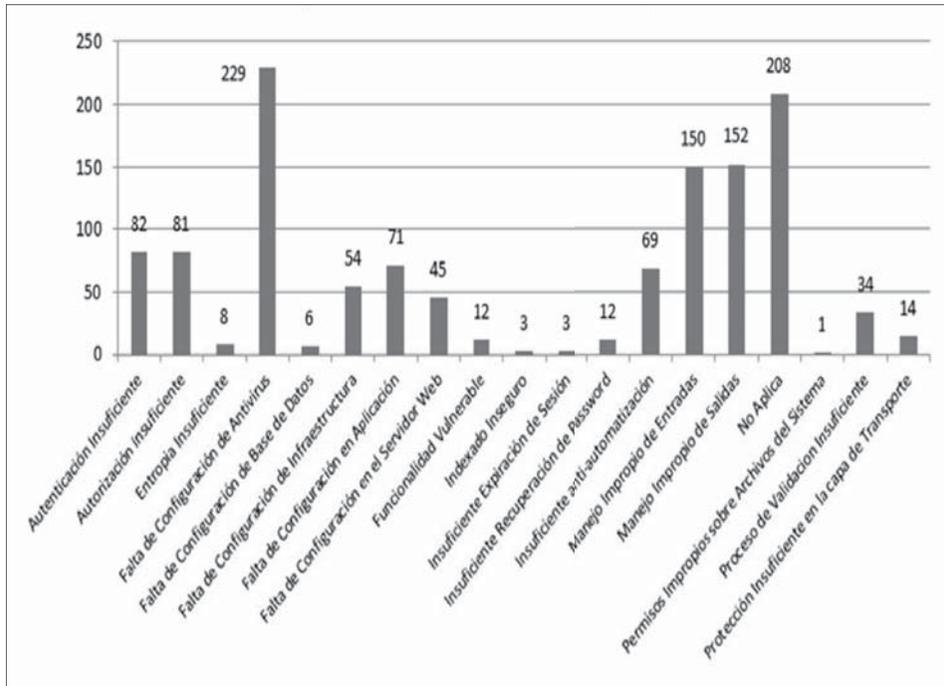


Figura 9. Distribución final de Valores Debilidad de la Aplicación

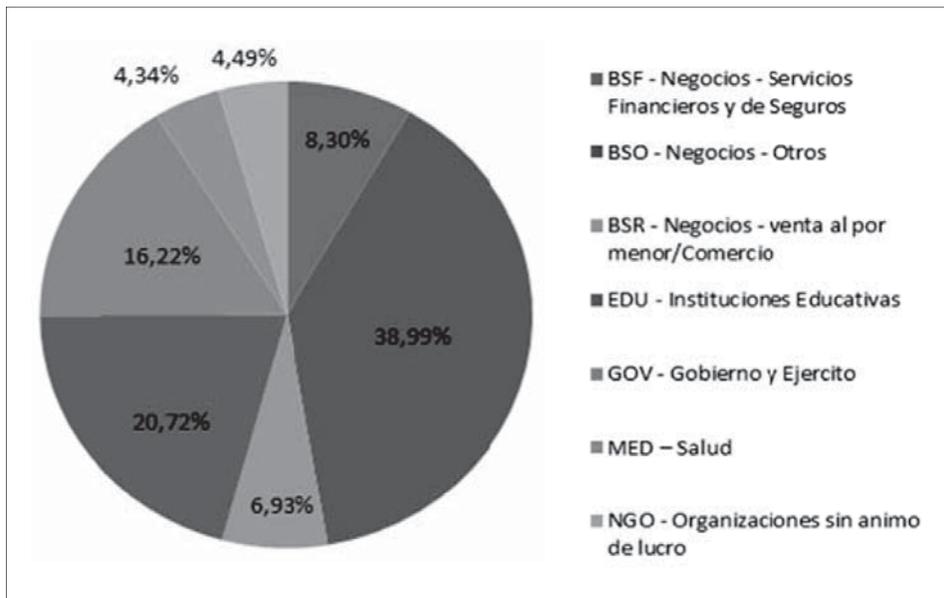


Figura 10. Distribución de incidentes por Tipo de Organización

7. Conclusiones

El reporte de incidentes de seguridad es una práctica que cada empresa debe tener en cuenta, esto con el fin de permitir rastrear y analizar los ataques y sus posibles relaciones, teniendo en cuenta que en Colombia las entidades financieras están reglamentadas por la circular 052 del 2007 de la Superintendencia Financiera [21] que fijó un estándar para la seguridad y calidad en el manejo de la información, enfocada a los datos de los clientes asegurando confidencialidad y disponibilidad. La taxonomía propuesta en este artículo se propone como solución para el reporte de los incidentes, pues permite determinar los perfiles de los mismos, con campos fáciles de diligenciar, evitando que los reportes sean descripciones ambiguas.

Según la norma ISO 27001[22] un sistema de gestión de seguridad (SGSI) debería implementarse en instituciones de cualquier naturaleza; en este sentido es de gran utilidad contar con un registro de los incidentes informáticos y un análisis estadístico de los mismos que permita encontrar las causas de un incidente que tengan como consecuencias inconformidades respecto a un conjunto de requerimientos.

De acuerdo a lo anterior, como trabajo futuro fruto de esta investigación, se propone la publicación del repositorio final obtenido en un sitio Web, además del desarrollo de una página que permita diligenciar los campos de la taxonomía propuesta para reportar nuevos incidentes de seguridad informática, con el fin de alimentar el repositorio con nuevos registros. Una vez esto se logre, el repositorio puede ser utilizado para investigaciones similares y/o enfocadas a la minería de datos, que permitan obtener conclusiones de gran valor para el área de seguridad informática, un tema que en la actualidad tiene una relevancia trascendental en el mundo.

Glosario

Concepto	Definición
Hacker	Un individuo el cual realiza acciones impulsado por reto, estatus, o la emoción de obtener acceso.
Criminales Profesionales	Ente el cual realiza acciones impulsado por obtener ganancias financieras para su propio bien.
Espías	Atacante el cual realiza acciones impulsado en obtener información la cual será utilizada para generar ganancia política.
Terroristas	Atacante el cual realiza acciones impulsado en generar miedo para obtener ganancia política.



Ladrón Corporativo	Individuo el cual realiza acciones impulsado en obtener información de la competencia corporativa para generar ganancia financiera.
Vándalos	Delincuente el cual realiza acciones impulsado en causar daño a una o varias entidades.
Agente Autónomo	Programa o fragmentos de programas los cuales operan independientemente del usuario. Algunos ejemplos son los virus o gusanos.
Ataque Físico	Robo físico o daño de un computador, red sus componentes o los sistemas soportados (tal como aire acondicionado, energía eléctrica, etc.)
Comando de usuario	Interacciones con una interfaz de entrada del sistema. Un ejemplo es ingresar comando de Unix a través de una conexión Telnet, o comandos un puerto SMTP.
Intercambio de Información	Un medio para obtener información desde otros atacantes o de personas siendo atacadas (llamada ingeniería Social).
Script o Programa	Indica el uso de programas diseñados específicamente para explotar vulnerabilidades o scripts que automatizan la ejecución de comando de usuario. Un ejemplo es un Shell Script para explotar un bug de un software, o un software de cracking para obtener Password.
Toolkit	Un paquete que contiene scripts, programas y/o agentes autónomos diseñado para facilitar la labor de atacantes inexpertos. Un ejemplo es el Toolkit llamado Rootkit.
Diseño	Fallo en la especificación de los requerimientos funcionales y/o no funcionales.
Implementación	Error en la codificación del software y/o falta de buenas prácticas de programación
Configuración	Falta de ajuste en los parámetros del software y/o hardware, tal como, credenciales con una contraseña por defecto, permisos de escritura globales para archivos, y/o contar con servicios vulnerables habilitados. También incluye la falta de actualización automática de software.

Flood	Acceder a un objetivo repetidamente con el propósito de sobrecargar la capacidad del objetivo.
Bypass	Evadir un proceso usando un método alternativo para acceder a un objetivo.
Spoof	Asumir la apariencia de una entidad diferente en una red de comunicaciones.
Cuenta	Entidad lógica compuesta por el acceso al dominio del usuario sobre un computador, red, o sistema; el cual es controlado acorde a un registro de información el cual contiene el nombre del usuario, contraseña y restricciones de uso.
Proceso	Conjunto formado por un programa ejecutable, los datos del programa, contador del programa, puntero del programa, y toda la otra información necesaria para ejecutar dicho programa.
Dato	Unidad lógica compuesta por representación de hechos, conceptos, o instrucciones, de una manera adecuada para comunicación, interpretación o procesamiento por humanos o por medios automáticos. Los datos puede encontrarse en forma de archivos en la memoria volátil o no volátil de un computador, o un dispositivo de almacenamiento, o en la forman de un dato en tránsito a través de un medio de transmisión.
Componente	Elemento físico que conforman un computador o red.
Computador	Entidad física compuesta por un dispositivo que consiste de uno o más componentes asociados, incluyendo unidades de procesamiento y unidades periféricas, que son controladas por programas almacenados y que pueden realizar procesos computacionales, que incluye operaciones aritméticas, u operaciones lógicas, sin intervención humana durante la ejecución.
Red	Entidad física y/o lógica compuesta por un grupo de host (computadores) interconectados o interrelacionados, elementos de switching, y ramas interconectadas.
Acceso incrementado	Obtención de privilegios de lectura, escritura y/o borrado no contemplados.
Difusión de la información	Divulgación de información privada, donde dicha información no es autorizada para la visualización a los usuarios que fue divulgada.



Corrupción de la información	Resultado de la alteración no autorizada de datos en un computador, sistema, o red.
Negación del Servicio	Desenlace de la degradación no intencional o bloqueo de un computador, recursos de red o un sistema.
Robo de recursos	Producto del uso no autorizado de un computador, sistema, o recursos de red.
Cambio de Estatus	Objetivo asociado a la necesidad de ser visto como un atacante habilidoso para violar la seguridad de cualquier sistema.
Daño	Finalidad asociada a la necesidad del atacante para ocasionar daño a cualquier sistema.
Ganancia Financiera	Intención asociada a la necesidad del atacante para obtener beneficios financieros.
Ganancia política	Búsqueda para obtener un mejor posicionamiento político, debido a la degradación verbal o visual de algún partido político, o grupos con pensamientos políticos.

Referencias

- [1] John D. Howard, Thomas A. Longstaff. "A Common Language for Computer Security Incidents". Sandia National Laboratories, Octubre 1998.
- [2] Carlos Carvajal, Diego Bayona, Análisis de Incidentes de Seguridad Informática Mediante Minería de Datos, para Modelado de Comportamiento y Reconocimiento de Patrones, Tesis para acceder al título de Ingeniería de Sistemas de la Universidad Francisco José de Caldas, Bogotá, Colombia, 2012
- [3] Web Application Security Consortium. "Web Hacking Incident DataBase", <http://projects.webappsec.org/w/page/13246995/Web-Hacking-Incident-Database>
- [4] Privacy Rights ClearingHouse, "Chronology of Data Breaches". <http://www.privacyrights.org/data-breach>
- [5] COL-CSIRT, Grupo de Investigación Universidad Distrital. <http://gemini.udistrital.edu.co/comunidad/grupos/arquisoft/colcsirt/>
- [6] CERT Brasil. "Centro de estudios de respuesta y tratamiento de incidentes de Seguridad de Brasil". <http://www.cert.br/>
- [7] INTECO España – Instituto Nacional del Tecnologías de la Comunicación. Centro de respuesta a incidentes de seguridad TIC. <http://cert.inteco.es>

- [8] "Hackers hit Tunisian websites" ALJAZEERA, Red de Noticias, 3 de Junio 2011, <http://www.aljazeera.com/news/africa/2011/01/201113111059792596.html>
- [9] "Secunia recovers from DNS redirection hack" The Register, Publicación Online de Tecnología, 26 de Noviembre de 2010. http://www.theregister.co.uk/2010/11/26/secunia_back_from_dns_hack/
- [10] "Cops: Hacker Posted Stolen X-rated Pics on Facebook" PCWorld, Magazín de Computación, 2 de Noviembre de 2010. http://www.pcworld.com/businesscenter/article/209584/cops_hacker_posted_stolen_xrated_pics_on_facebook.html
- [11] COMPUTERWORLD, Magazín de Computación. http://www.computerworld.com/s/article/9149218/Bank_sues_victim_of_800_000_cybertheft
- [12] "Bank sues victim of \$800,000 cybertheft" The Washington Post, Periódico de la ciudad de Washington, 26 de Enero de 2010. <http://www.washingtonpost.com/wp-dyn/content/article/2010/03/18/AR2010031805464.html>
- [13] "Hackers Take Over BP Twitter Feed" FOX NEWS, Red de noticias, 27 de Mayo de 2010. <http://www.foxnews.com/tech/2010/05/27/hackers-bp-twitter-feed/>
- [14] Jiawei Han, Micheline Kamber. "Data Mining: Concepts and Techniques". Morgan Kaufmann Publishers 2000, Capitulo 3 "Data Preprocessing".
- [15] Roderick J. A. little, Donald B. Rubin. "Statistical Analysis with Missing Data". Capitulo 6 "Theory of Inference Based on the Likelihood Function".
- [16] Dorian Pyle. "Data Preparation for Data Mining". Capitulo 8 "Replacing Missing and Empty Values". http://www.temida.si/~bojan/MPS/materials/Data_preparation_for_data_mining.pdf
- [17] Jaisheel Mistry. "Estimating Missing Data and Determining the Confidence of the Estimate Data" Seventh International Conference on Machine Learning and Applications, 2008.
- [18] Ludmila Himmelspach. "Clustering Approaches for Data with Missing Values: Comparison and Evaluation", Institute of Computer Science Heinrich-Heine-Universität Düsseldorf, 2010.
- [19] Jiawei Han, Micheline Kamber. "Data Mining: Concepts and Techniques". Morgan Kaufmann Publishers 2000, Capitulo 8 "Cluster Analysis".
- [20] Xiao-Bai Li . "A Bayesian Approach for Estimating and Replacing Missing Categorical Data". ACM Journal of Data and Information Quality, Vol. 1, No. 1, Article 3, Junio 2009.
- [21] "ABC de la Circular Externa 052 DE 2007", Comunicado de Prensa, Superintendencia Financiera de Colombia, Noviembre 7 de 2007. <http://www.superfinanciera.gov.co/>
- [22] "ISO/IEC 27001 : 2005 Tecnología de la Información- Técnicas de seguridad- Sistemas de gestión de seguridad- Requerimientos ". Estándar Internacional, Octubre 2005. <http://www.iso.org/iso/home.html>



Carlos Javier Carvajal Montealegre

Es Ingeniero de Sistemas de la Universidad Distrital Francisco José de Caldas de Bogotá, Colombia. Participó en el grupo de trabajo ACM UD-GIIA (capítulo estudiantil de la Association for Computing Machinery de la Universidad Distrital Grupo de Interés en Inteligencia Artificial) como líder del tema de algoritmos genéticos para prototipo de agente autónomo (2008). Participó en el grupo de investigación GIIIRA-UD de la Universidad Distrital en el tema Factores Humanos y Multimedia (2007-2008). Actualmente se desempeña como Analista Programador y Líder Técnico de Proyectos de Software. e-mail: lng.carlosj@gmail.com

Diego Nicolás Bayona Sastoque

Es Ingeniero de Sistemas de la Universidad Distrital Francisco José de Caldas de Bogotá, Colombia. Participó en el grupo de trabajo ACM UD-GIIA (capítulo estudiantil de la Association for Computing Machinery de la Universidad Distrital Grupo de Interés en Inteligencia Artificial) como líder del tema de algoritmos genéticos para prototipo de agente autónomo. 2008. Actualmente se desempeña como Services Manager en Spring Mobile Solutions S.A.S. e-mail: nicolas.bayona@gmail.com

Zulima Ortiz Bayona

Es Matemática de la Universidad Nacional, Bogotá, Colombia. Obtuvo su título de Maestría en Matemáticas en la Universidad Nacional de Colombia. Es especialista en Teleinformática de la Universidad Distrital Francisco José de Caldas. Es integrante del Grupo de Investigación Arquisoft Rama de Seguridad Informática de la Universidad Distrital Francisco José de Caldas. Profesora Universidad Nacional. Actualmente se desempeña como profesora en el área de Matemáticas en la Universidad Distrital Francisco José de Caldas en Bogotá, Colombia. e-mail: zortiz@udistrital.edu.co

Identificación de relaciones entre los nodos de una red social

Identification of node relationships in a social network

Mónica Andrea Niño Barón

Estudiante Maestría
Universidad Distrital
Grupo GESDATOS
ninoandrea75@hotmail.com

Sonia Ordóñez Salinas

Docente
Universidad Distrital
Directora Grupo GESDATOS
sordonez@udistrital.edu.co



Resumen

El presente artículo realiza una revisión del tema, representación y clasificación de relaciones de pertenencia entre los nodos de una red social. Para ello, se abordan aspectos sobre Procesamiento de Lenguaje Natural, Minería de Texto, Recuperación de Información y Entidades Nombradas. Se hace una descripción de cada una de ellas y se referencian y discuten trabajos académicos destacados que se han desarrollado en dicho tema.

Palabras clave: Red social, nodos, pertenencia, Procesamiento de Lenguaje Natural, entidades nombradas.

Abstract

In this paper a review is conducted about representation and classification of membership among nodes belonging to a social network. For this purpose, topics such as Natural Language Processing, Text Mining, Information Retrieval and Named Entities are considered description and survey of outstanding approaches is carry out in each topic.

Key words: social network, nodes, membership, Natural Language Processing, named entities.

1. Introducción

Los servicios de redes sociales involucran a millones de usuarios en línea. Aunque el concepto de una red social se ha establecido desde comienzos del siglo 20 [1], tan solo en los últimos años se ha comenzado a masificar su uso en internet gracias a servicios como *Facebook*, *Twitter* o *LinkedIn*. Su número de visitantes únicos en Latinoamérica pasó de 53.248.000 en 2008 [2] a 139.000.000 en 2013[3, 4].

La información que puede extraerse de dichos servicios presenta características importantes para los investigadores de diferentes áreas, y por

Fecha recibido: abr. 6/2013
Fecha modificado: jun. 5/2013
Fecha aceptado: jun. 14/2013



lo tanto, su crecimiento también permite contar con repositorios de información cada vez más grandes como los de las universidades de Arizona [5] y Stanford [6], el proyecto Ancora [7], el American National Corpus [8] y Molinero en Español [9], entre otros. La cantidad de información hace que la calidad de la misma pueda disminuirse, y en ocasiones, los márgenes de error no sean los adecuados.

La resolución de preguntas en estas grandes colecciones concernientes a ubicar una persona como miembro de una comunidad, debe considerar por un lado el uso de técnicas para reconocimiento de entidades nombradas (todo término o palabra que corresponde a un nombre propio que identifica una entidad del mundo real y no se encuentra en los diccionarios [10]), y por el otro, técnicas para establecer si dichas entidades pertenecen a la comunidad o más genéricamente a un dominio de investigación, o descartarlas en su defecto. A esta temáticamente se le podría denominar *búsqueda de relaciones de pertenencia*.

En el presente artículo se presentan los avances en dicho tema. Para ello, se ha dividido el trabajo en tres secciones: En la primera sección se explicará el marco de las relaciones de pertenencia y las entidades nombradas dentro del Procesamiento de Lenguaje Natural (PLN). En una segunda sección, se presenta el estado del arte sobre la identificación de relaciones de pertenencia en redes sociales, el modelo matemático con que se estudia el tema y las métricas propuestas. Y en la última sección se presentan las herramientas para el apoyo en tareas de PLN. El artículo finaliza con las conclusiones del estudio.

2. Procesamiento de Leguaje Natural

El término se utiliza para describir la función que cumplen los sistemas de hardware y de software dedicados a analizar el lenguaje hablado o escrito. [11].

Dentro de las técnicas asociadas al procesamiento de lenguaje natural se encuentran la minería de texto y la extracción de texto e información, que se definen a continuación:

2.1 Minería de texto

La minería de texto es el proceso de extraer patrones interesantes a partir de grandes colecciones de textos para descubrir conocimiento [12]. Es también el descubrimiento de reglas de asociación importantes dentro de un corpus de texto [13]. Usualmente se utilizan en corpus de gran tamaño, como es el caso de la información obtenida de servicios de redes sociales, u otros servicios soportados en internet.

En relación al tema de estudio, aparece en el año 2004, en las memorias de IEEE/WIC/ACM International Conference, un estudio sobre inferencias en conversaciones de

chat [14] utilizando técnicas de minería de texto, con el objeto de determinar el tema principal en una sala de chat. En dicho estudio se observa un análisis completo de la situación, incluyendo manejo de ruido, expresiones concisas y dinámicas, texto cambiante, etc. Los autores utilizan el modelo de clasificación de texto SVM (Máquina de Vectores de Soporte [15]) dejando planteada la posibilidad de manejar los participantes del chat como una red social.

Posteriormente en el año 2006, Feldman publica un libro [16] sobre minería de texto con una recopilación de trabajos, técnicas y herramientas para minería de texto. Sin embargo, en él todavía no se contemplaban aplicaciones a dominios de redes sociales.

A partir del 2008, con el auge de las redes sociales en Internet como *Facebook* y *Twitter*, se proponen estudios sobre técnicas de minería de texto para extraer información en dichas redes. En la mayoría de ellos se acude a una copia de los datos fuera de línea (descargados de diferentes servicios de redes sociales), se procesan y analizan dichos datos, y en ocasiones se ofrecen herramientas para su visualización. Hay un patrón en tales trabajos y es la aceptación de las técnicas y métricas propuestas por X. Chunyan [17] quince años atrás, sobre el manejo de redes sociales con base en grafos. Hui [18] por ejemplo, propone utilizar técnicas de minería de texto para extraer grupos con intereses comunes dentro de una red de blogs. Fard [19] utiliza el mismo concepto para descubrir grupos criminales. Nieto [20] presenta un trabajo similar en español (entendiendo que cada lenguaje tiene sus propias reglas y los trabajos difieren en metodología, algoritmos y enfoque). Todos ellos comienzan a utilizar datos expuestos en los servicios de redes sociales, con las técnicas de Wasserman, pero incorporando mejoras para cada lenguaje y escenario.

En el año 2010, se presenta un libro con casos de estudio sobre minería aplicada a redes sociales [21]. El trabajo se basa en las memorias y artículos presentados en el MSN-DS 2009 de Atenas, Grecia (*Minning Social Networks for Decision Supports Work-shop*), considerado el primer evento internacional sobre minería de texto aplicada a las redes sociales. El libro muestra el uso del análisis de redes sociales para el descubrimiento de patrones sobre ellas mismas.

A comienzos de 2010 aparecen trabajos con técnicas variadas [21, 22]. Wu [23] por ejemplo propone un método de detección de patrones de texto a través de la utilización de un “multígrafo dirigido”, previamente enunciado por Feldman [16]; adicionalmente se desarrolla un algoritmo basado en la distancia de las palabras clave; el método se aplica en la detección de plagio de documentos y correos electrónicos fraudulentos escritos por la misma persona, obteniendo resultados destacados.

Recientemente, Rusell [24] planteó en 2011 un trabajo para extraer información de redes sociales, basado en etiquetas de HTML5, con javascript; dicho trabajo fue validado y extendido por M. Fire [25] en 2013, soportado con tecnología Ajax.



2.1.1 Preprocesamiento de texto

El preprocesamiento de texto es una de las tareas iniciales y primordiales en las actividades de procesamiento del lenguaje natural; sus actividades más importantes son: eliminación del ruido o eliminación de palabras no relevantes [26], identificación de la raíz de las palabras (llamado “lematización” o “stemming”) y la ruptura del texto en párrafos y palabras, lo cual es llamado tokenización [27]. Manning [28, 29] sintetiza diferentes técnicas para cada una de dichas actividades.

El preprocesamiento de texto suele cambiar según el lenguaje. En algunos de ellos como el inglés, las reglas para conjugación de verbos son simples pero los verbos en ocasiones pueden actuar como sustantivos; en español es común evitar el sujeto, o utilizar verbos auxiliares, las excepciones a las reglas del lenguajes son muy extensas y muchas palabras suelen ser nombres propios, verbos y sujetos a la vez. En alemán, la mayoría de las palabras son a su vez compuestas por otras, algo que no es usual en otros lenguajes. Otros de ellos utilizan incluso códigos diferentes, lo cual hace a cada uno especial. Igualmente, el dominio del problema (el campo al cual se aplicará la técnica) puede variar. Por ejemplo, el vocabulario especializado de una u otra profesión hacen variar la técnica entre sí.

Hay pocas coincidencias entre trabajos, pero es posible determinar algunas tendencias. En cuanto a lematización en español e inglés se suele utilizar el algoritmo de Porter, basado en las reglas comunes del lenguaje y condicionales simples. Para la tokenización en español e inglés se suele separar por palabras o por frases sencillas, y para la lematización se suele utilizar tesauros y ontologías.

2.1.2 Entidades nombradas (NE)

En 2005 hubo una serie de trabajos como el de Peng [30] para extraer entidades nombradas en un corpus grande de texto, el de Tomita [31], que presenta un algoritmo para reconocimiento de entidades nombradas con un sistema jerárquico de palabras, el trabajo de Diamantaras [32] que utiliza un sistema de clasificación “linear binario” para determinar entidades nombradas en un texto, y el de Pu-Jen [33], que propone un método no supervisado para detectar entidades nombradas basado en el análisis del documento completo.

En 2008, Todorovic [34] desarrolla un algoritmo basado en cadenas ocultas de Markov para el reconocimiento de entidades nombradas de tipo persona, localización y organización; en 2009, Jianhan [35], se presenta un trabajo sobre un *framework* para reconocimiento de entidades nombradas escalable diseñado para Web.

Appice en el 2010 [36] publica un trabajo que abarca extracción de entidades nombradas en literatura biomédica. Es importante resaltarlo debido a que cada posible diccionario técnico representa diferentes retos en un determinado algoritmo.

En el 2011 Jung [37] propuso un método para la identificación de entidades nombradas en microtextos, partiendo de la agrupación de los mismos y ya que la red social se encuentra automáticamente construida, los vínculos sociales son utilizados como base para extraer de una mejor manera las entidades nombradas. También Bollegala [38] realiza estudios sobre la identificación de entidades nombradas en microtextos o textos breves que se publican en redes sociales como Twitter y Facebook. Otros estudios como el de Tkatchenko [39] propone un enfoque semi-supervisado para la construcción de conjuntos de entrenamiento para la clasificación de entidades nombradas. Para su desarrollo se usó una taxonomía de entidades nombradas llamada BBN [40], un umbral de al menos 40 artículos de Wikipedia, y un subconjunto de las 400 palabras en minúscula más frecuentes, del corpus Reuters.

Otros trabajos relacionados con entidades nombradas se refieren a la búsqueda e identificación de seudónimos de personas en la Web [38], mediante la extracción de patrones léxicos y su posterior clasificación.

2.1.3 Extracción y representación de relaciones de asociación.

La tesis doctoral de Jiménez Ruiz [41] desarrolla un estudio sobre un modelo formal para la extracción de reglas difusas, plantean varias propuestas para la extracción del conocimiento de bases de datos, utilizando teorías de subconjuntos difusos. Proponen nuevas reglas para la descripción de la relación entre dos conjuntos de ítems; por último plantean un procedimiento para la extracción de reglas de asociación.

Existe otro estudio realizado por Uday Kiran [42], sobre la extracción de reglas de asociación soportado en elementos diferentes o “raros” (el autor se basa en la teoría de que en cualquier dominio de problema los casos comunes son los más sencillos de analizar pero aportan menor precisión, mientras que las excepciones inciden más en la precisión de un caso de uso, es decir, que se le debe asignar mayor ponderación a las palabras o frases menos comunes en una colección [24, 43]). El autor asigna un peso a cada palabra o frase mediante su frecuencia en la colección; calcula la cantidad de veces que aparece cada palabra en su documento y multiplica dicho valor por una frecuencia inversa, que incluya el resto de palabras de la colección para darle mayor relevancia a las palabras cuando sean más relevantes por fuera del documento. Esta técnica es conocida como TF-IDF [44, 45].

2.1.4 Clasificación de información

Dentro de los algoritmos de clasificación supervisada, para texto, se encuentran las SVM (Máquinas de Vectores de Soporte, método lineal basado en la maximización de la separación entre dos clases distintas de vectores proyectados en espacios de mayor dimensionalidad [15]), ANN (Redes Neuronales Artificiales [46], que son algoritmos que simulan



el comportamiento neuronal, en los cuales aparecen estructuras de información llamada nodos o neuronas, enlazadas entre sí por medio conexiones), Regresión Logística [47] (técnica en la cual los valores de clasificación se seleccionan por aproximación a partir de una serie de funciones dependientes del modelo), Naive-Bayes (basado en inferencias probabilísticas de la teoría de Bayes) [16], KNN (Clasificación por similitud con los vecinos más cercanos [48]) y árboles de decisión [15] (construidos con técnicas de teoría de la información y análisis de entropía de los datos). Cada uno de ellos presenta características diferentes de rendimiento según su uso [15].

Por otra parte se encuentran los grafos conceptuales [49] (como formalismo estructurado y clasificado como estructura conceptual) permiten representar conocimiento a través de aristas entre dos tipos de nodos: conceptos y relaciones, palabras y símbolos propios de la lógica matemática y del lenguaje natural.

2.1.5 Relaciones entre entidades nombradas

Existen diferentes tipos de relaciones entre miembros de un mismo conjunto y entre los conjuntos en sí mismos. Cuando se tiene una muestra o corpus con una serie de entidades nombradas, es deseable determinar cuál (o cuáles) debe actuar como conjuntos, y cuáles deben actuar como objetos pertenecientes a dicho conjunto.

Existen trabajos al respecto, como el desarrollado en 2004 por Hasegawa [50], que definía cinco pasos en el proceso de asociación entre entidades nombradas: marcar las entidades nombradas (en inglés se utiliza el término *tagging*), identificar casos recurrentes, medir la cantidad de similitudes dentro del contexto, hacer conjuntos de pares y etiquetar cada conjunto de pares; en 2005 Cheng [33] utiliza el mismo concepto de pares de entidades nombradas, pero ponderando cada par según la similitud en todo el corpus (asignando un valor de cero cuando no hay suficiente similitud y un valor alto en la medida en que dicho par de entidades nombradas fuese más encontrado dentro de los documentos). Al obtener un alto grado de precisión, es un modelo que se utiliza actualmente. En 2007, Hirano [51] propone adicionar un mecanismo de aprendizaje supervisado al proceso, el cual mejora en un 4.4% la precisión. En 2011, Tkatchenko [39] propone utilizar un clasificador con aprendizaje semi-supervisado basado en SVM para establecer relaciones entre entidades nombradas dentro de *Wikipedia*. El trabajo ofrece niveles de precisión cercanos a 1 (100%) al aplicar el clasificador sobre 18 clases, lo cual es un resultado destacable.

2.2 Recuperación de información

A nivel internacional, hay trabajos, en el que se hace estudios comparativos entre métodos para extracción de atributos de personas en la Web (en inglés) [52]. En este estudio, se realizan entre otros experimentos, la recuperación “superficial”, en el que se extraen etiquetas de marcado HTML y se intenta obtener entidades nombradas desde allí; se

compara contra una búsqueda en profundidad. La comparación entre los métodos se hace en términos de las medidas de desempeño *Recall* y *Precision*.

Douglas y Gong [53], plantean una técnica de desambiguación de nombres por medio de *Hierarchical Clustering* (Agrupación Jerárquica), que utiliza la combinación de algunos métodos planteados por Artiles [54], Manning [28], Schütze [29] y Elmacioglu [55], con la técnica de aprendizaje SVM (Cristianini [56]).

En 2010 se reportan casos de estudio [23, 57, 58] sobre la recuperación de información en redes sociales, mezclando técnicas ya conocidas pero sin planteamientos especiales. En 2011 continuó la misma tendencia [59, 60].

En cuanto a aplicaciones en español, existen avances en Extracción de Información basado en técnicas específicas, como el planteado por Roperio [61] (Método general de Extracción de información basado en el uso de Lógica Borrosa: aplicación en portales Web).

3. Identificación de relaciones en redes sociales

Jamail [62] define una red social como una estructura social entre actores en su mayoría personas u organizaciones. Dicha estructura se basa en vínculos sociales, económicos, y de cualquier otra índole.

3.1 Modelo computacional

En el desarrollo de modelos computacionales de redes sociales, recientes estudios se han enfocado en diseño y elaboración de servicios que permitan enriquecer con diferentes componentes tecnológicos, como celulares, GPS, etc. El proyecto realizado por Jung-Tae [63], pretende ofrecer formas más inteligentes y activas para intercambios en el sistema de información. Está compuesto por un integrador de servicios sociales, ubicación y una ontología social, ayudando a la interpretación semántica de los usuarios y su información de manera casi automática. Como herramientas para la representación de la ontología se utilizó XFN (*XHTML Friends Network*) y OWL (*Ontology Web Language*) para la modificación de relaciones. También en el 2010, se publica un libro escrito por Furht [57] en el que se observan las tecnologías y aplicaciones para el manejo de redes sociales, se presentan casos de estudio y se estudian las tendencias a nivel operativo de las diferentes herramientas.

3.2 Métricas para establecer relaciones en una red social

En 2003, James Moody de la Universidad de Ohio [64], comienza a analizar las redes sociales desde el punto de vista jerárquico, manejando matemáticamente el tema como árboles. En este estudio, más social y antropológico, se establece además el concepto de



cohesión estructural, según la capacidad de extraer miembros del grupo sin que el grupo se divida. En el estudio se presenta un algoritmo simple para determinar dicha medida.

Ulrik Bandes [65] habla de herramientas de análisis para fenómenos sociales y establece la centralidad como una de las herramientas, basada en la ruta más corta definida en teoría de grafos. Newman [66] establece la centralidad (intermediación, cercanía o grado) que se ha utilizado para evaluar las redes sociales desde un modelo matemático. Define, basado en la teoría de grafos, el nodo más importante (central) según la suma de sus conexiones.

Se ha avanzado poco en buscar nuevas métricas o algoritmos para perfeccionar las anteriores. Las teorías de grafos han sido aceptadas como herramientas de análisis, y los algoritmos que se utilizan tienen más de 20 años de utilización y aceptación. Por tanto, se entienden como aceptadas las siguientes métricas para establecer y cuantificar relaciones en una red social, entendida como un grafo:

Centralidad (intermediación, cercanía, grado): La centralidad es un atributo estructural de los nodos en una red. Se trata de un valor asignado al nodo debido a su posición estructural en la red. Por ejemplo en un grafo en forma de estrella, el nodo central ocupa un valor máximo de centralidad, mientras que los nodos de las puntas ocupan un valor de centralidad inferior.[67]

Conectividad (puente). Un “punto de quiebre” de un grafo es un conjunto de aristas que, al ser removidas, dejan el grafo inconexo. Un puente es un punto de quiebre de una arista [68].

Coefficiente de agrupamiento: El coeficiente de agrupamiento (*clustering coefficient*) de un vértice en un grafo es la medida para cuantificar el nivel de interconexión de dicho nodo con sus vecinos [69]

Densidad: La densidad de un grafo define la cantidad de posibles aristas (relaciones) utilizadas en él. Técnicamente, un grafo puede tener una densidad entre 0 (cuando ningún nodo está conectado) y 1 (cuando todos los nodos están conectados con todos los demás nodos). Es decir, que un grafo denso es un grafo en el que el número de aristas está cercano al número máximo de aristas. Lo opuesto, un grafo con solo algunas aristas, es un grafo disperso [10].

3.3 Estructuras para la representación de redes sociales

El estudio de las redes sociales ha dado origen al diseño de modelos que permitan representar de alguna manera el conocimiento y las relaciones entre ellas.

Para la representación de las redes sociales generalmente se utilizan métodos formales (matemáticas y grafos), ya que permiten representar las descripciones de las redes de for-

ma compacta y sistemática, así mismo utilizar herramientas informáticas para el análisis de la red de datos [62]. Igualmente, para la representación de las relaciones en las redes sociales, las herramientas más conocidas son los grafos y las matrices.

3.3.1 Grafos

Junhua [70] define un grafo como un conjunto de nodos (actores) y un conjunto de líneas (relaciones) que conectan los nodos.

Según Jamali [62] el uso de los grafos en redes sociales pequeñas es muy útil, mientras que en redes grandes su lectura e interpretación se hace compleja. Ramanathan [71] considera que los grafos son útiles para la representación binaria de las relaciones entre nodos, pero no cuando existe propiedades de grupo diferentes.

3.3.2 Estructuras vectoriales y modelos estadísticos

Shoaib [72] desarrolló una ontología que representa la personalización en sitio de las redes sociales, permite el almacenamiento de usuarios, grupos, mensajes, perfil de los acontecimientos en formato de máquina para poder procesarlos después. Utiliza fuentes externas a la red social para la personalización del sitio de los usuarios de la red y cargue de perfiles, los cuales están almacenados de manera semántica. Este estudio se ha hecho solo para la academia.

3.3.3 Herramientas para el apoyo en tareas del procesamiento de Lenguaje natural

Cuando se realizan tareas de procesamiento de lenguaje natural, se dedica un tiempo considerable en la selección y evaluación de herramientas para la realización de dichas tareas. Lobur [73] realizó una evaluación de *Natural Language Toolkit* y su utilización en el campo educativo en el curso de lingüística computacional de la *Lviv Polytechnic National University*, concluyendo que una alta proporción de la muestra de estudiantes, que no contaban con conocimientos en programación, adquirieron habilidades en el Procesamiento de Lenguaje Natural.

En la Universidad Nacional de Colombia se desarrolló una guía metodológica para la selección de técnicas de depuración de datos [74]. Dicha guía está justificada sobre los siguientes postulados: ninguna métrica es adaptable a todos los conjuntos de datos, es poco probable que se resuelva pronto la pregunta de cuál de los métodos, debe utilizarse para una determinada tarea de depuración. La tarea de depuración de datos, es altamente dependiente de los datos y no es evidente que exista una técnica que domine a todas las demás en todos los conjuntos de datos. Y define dentro de la metodología técnicas para detección de duplicados, corrección de valores faltantes y detección de valores atípicos.



4. Conclusiones

Para el descubrimiento de reglas de asociación en un texto, se deben cumplir con las siguientes etapas: preprocesamiento de texto, identificación de entidades nombradas, representación de relaciones de asociación, clasificación de la información y establecimiento de relaciones de pertenencia entre las entidades nombradas. Así mismo si se desean identificar relaciones entre los miembros o nodos de las redes sociales, se requiere que la información o el texto a analizar este estructurado o tenga una representación formal. Esto puede ser una fuente de trabajos futuros, ya que se hace importante explorar técnicas en cualquier etapa de las mencionadas anteriormente, enmarcándolo en un idioma determinado y en un dominio de problema específico.

En lo que respecta a la representación de relaciones en redes sociales, se puede realizar utilizando diferentes estructuras, como son: grafos, estructuras vectoriales y modelos estadísticos, aunque se encuentra abierta la posibilidad de representar relaciones con diferentes técnicas.

Son muchos los trabajos encontrados en el ámbito del procesamiento de lenguaje natural aplicado a las redes sociales; es un tema que está en vigor y amerita profundizarlo. Así mismo en la práctica es posible encontrar aplicabilidad o cualquier avance en ramas como el mercadeo (inteligencia de negocios, fuerza de ventas, tendencias y preferencias), en educación (metodologías basadas en tendencias, establecimiento de nuevas técnicas) y seguridad, este último siendo el más estudiado hasta ahora.

Referencias bibliográficas

- [1] G. Simmel, "Conflict," in *Conflict and the Web of Group Affiliations*, T. F. Press, Ed., ed Glencoe, IL: The Free Press, 1908.
- [2] A. Lipsman, "Social Networking Explodes Worldwide as Sites Increase their Focus on Cultural Relevance," comScore2008.
- [3] A. L. Zain, "Futuro digital Latinoamérica 2013: El estado actual de la industria digital y las tendencias que están modelando el futuro," ComScore2013.
- [4] J. Seguic, "El Crecimiento de Redes Sociales en América Latina: La Influencia de Los Medios Sociales en el Escenario Digital de América Latina. Septiembre 2011.," ComScore2011.
- [5] Z. R. a. L. H. (2009, 04/07/2013). *Social Computing Data Repository at {ASU}*. Available: <http://socialcomputing.asu.edu>
- [6] J. McAuley and J. Leskovec, "Learning to discover social circles in ego networks," in *Advances in Neural Information Processing Systems 25*, 2012, pp. 548-556.
- [7] C.-C. d. L. i. Computació. (04/07/2013). *AnCorra*. Available: <http://clic.ub.edu/corpus/es>
- [8] N. Ide and C. Macleod, "The american national corpus: A standardized resource of american english," in *Proceedings of Corpus Linguistics 2001*, 2001.

- [9] M. A. Molinero, B. Sagot, and L. Nicolas, "A morphological and syntactic wide-coverage lexicon for Spanish: The Leffe," in *Proceedings of the 7th conference on Recent Advances in Natural Language Processing (RANLP 2009)*, 2009.
- [10] B. Preiss, *Data Structures and Algorithms with Object-Oriented Design Patterns in C++*: John Wiley & Sons, 1998.
- [11] P. Jackson and I. Moulinier, *Natural language processing for online applications: text retrieval, extraction and categorization*: John Benjamins Pub., 2007.
- [12] A. H. Tan, "Text Mining: promises and challenges," *South East Asia Regional Computer Confederation, Singapore*, 1999.
- [13] M. Delgado, N. Marin, D. Sanchez, and M. A. Vila, "Fuzzy association rules: general model and applications," *Fuzzy Systems, IEEE Transactions on*, vol. 11, pp. 214-225, 2003.
- [14] V. H. Tuulos and H. Tirri, "Combining Topic Models and Social Networks for Chat Data Mining," in *Web Intelligence, 2004. WI 2004. Proceedings. IEEE/WIC/ACM International Conference on*, 2004, pp. 206-213.
- [15] R. Caruana and A. Niculescu-Mizil, "An empirical comparison of supervised learning algorithms," presented at the Proceedings of the 23rd international conference on Machine learning, Pittsburgh, Pennsylvania, 2006.
- [16] R. S. Feldman, J., *The Text Mining Handbook*. New York: Cambridge University Press, 2006.
- [17] X. Chunyan, "Human-machine interface evaluation method based on grey interval relation membership degree," in *Advanced Computer Control (ICACC), 2010 2nd International Conference on*, 2010, pp. 288-291.
- [18] W. Hui-Ju, I. H. Ting, and W. Kai-Yu, "Combining Social Network Analysis and Web Mining Techniques to Discover Interest Groups in the Blogspace," in *Innovative Computing, Information and Control (ICICIC), 2009 Fourth International Conference on*, 2009, pp. 1180-1183.
- [19] A. M. Fard and M. Ester, "Collaborative Mining in Multiple Social Networks Data for Criminal Group Discovery," in *Computational Science and Engineering, 2009. CSE '09. International Conference on*, 2009, pp. 582-587.
- [20] A. Nieto Martín and M. Maroto Calatayud, "Redes sociales en internet y" data mining" en la prospección e investigación de comportamientos delictivos," in *Derecho y redes sociales*, 2010, pp. 207-258.
- [21] I.-H. Ting, H.-J. Wu, and T.-H. Ho, *Mining and Analyzing Social Networks*: Springer Publishing Company, Incorporated, 2010.
- [22] G. Xu, Y. Zhang, and L. Li, *Web Mining and Social Networking: Techniques and Applications*: Springer-Verlag New York, Inc., 2010.
- [23] Q. Wu, E. Fuller, and C.-Q. Zhang, "Graph Model for Pattern Recognition in Text Mining and Analyzing Social Networks." vol. 288, I. H. Ting, H.-J. Wu, and T.-H. Ho, Eds., ed: Springer Berlin / Heidelberg, 2010, pp. 1-20.
- [24] M. Russell. (2011). *Mining the Social Web: Analyzing Data from Facebook, Twitter, LinkedIn, and Other Social Media Sites*. Available: http://books.google.com.co/books?id=SYM1lrQd-rdsC&printsec=frontcover&dq=Mining+the+Social+Web:+Analyzing+Data+from+Facebook,+Twitter,+LinkedIn&hl=es&ei=S-FRTo2ZHMHR0gGWl82sBw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC0Q6AEwAA#v=onepage&q&f=false



- [25] M. Fire, R. Puzis, and Y. Elovici, "Organization Mining Using Online Social Networks," *arXiv preprint arXiv:1303.3741*, 2013.
- [26] Facebook. (2012, 20/02/2012). *Facebook developers*. Available: <http://developers.facebook.com/>
- [27] Twitter. (2012, 20/02/2012). *Twitter developers*. Available: <https://dev.twitter.com/>
- [28] C. Manning, Raghavan, P., Schütze, H, *An introduction to information retrieval*, 2009.
- [29] C. D. Manning and H. Schütze, *Foundations of statistical natural language processing*: MIT Press, 1999.
- [30] L. Peng, W. Xiao Long, G. Yi, and Z. Yu Ming, "Extracting answers to natural language questions from large-scale corpus," in *Natural Language Processing and Knowledge Engineering, 2005. IEEE NLP-KE '05. Proceedings of 2005 IEEE International Conference on*, 2005, pp. 690-694.
- [31] T. Tomita, Y. Okimoto, H. Yamamoto, and Y. Sagisaka, "Speech recognition of a named entity," in *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP '05). IEEE International Conference on*, 2005, pp. 1057-1060.
- [32] K. I. Diamantaras, I. Michailidis, and S. Vasiliadis, "A Very Fast and Efficient Linear Classification Algorithm," in *Machine Learning for Signal Processing, 2005 IEEE Workshop on*, 2005, pp. 93-98.
- [33] C. Pu-Jen, C. Hsin-Chen, P. Yi-Cheng, and C. Lee-Feng, "Annotating text segments in documents for search," in *Web Intelligence, 2005. Proceedings. The 2005 IEEE/WIC/ACM International Conference on*, 2005, pp. 317-320.
- [34] B. T. Todorovic, S. R. Rancic, I. M. Markovic, E. H. Mulalic, and V. M. Ilic, "Named entity recognition and classification using context Hidden Markov Model," in *Neural Network Applications in Electrical Engineering, 2008. NEUREL 2008. 9th Symposium on*, 2008, pp. 43-46.
- [35] Z. Jianhan, "An adaptive approach for web scale named entity recognition," in *Web Society, 2009. SWS '09. 1st IEEE Symposium on*, 2009, pp. 41-46.
- [36] A. Appice, M. Ceci, and C. Loglisci, "Discovering Informative Syntactic Relationships between Named Entities in Biomedical Literature," in *Advances in Databases Knowledge and Data Applications (DBKDA), 2010 Second International Conference on*, 2010, pp. 120-125.
- [37] J. J. Jung, "Towards Named Entity Recognition Method for Microtexts in Online Social Networks: A Case Study of Twitter," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, 2011, pp. 563-564.
- [38] D. Bollegala, Y. Matsuo, and M. Ishizuka, "Automatic Discovery of Personal Name Aliases from the Web," *Knowledge and Data Engineering, IEEE Transactions on*, vol. 23, pp. 831-844, 2011.
- [39] M. Tkatchenko, A. Ulanov, and A. Simanovsky, "Classifying Wikipedia entities into fine-grained classes," in *Data Engineering Workshops (ICDEW), 2011 IEEE 27th International Conference on*, 2011, pp. 212-217.
- [40] J. Y. Zhang, Y. (2003) Robustness of Regularized Linear Classification Methods in Text Categorization. *Proceedings of SIGIR-03, 26th ACM International Conference on Research and Development in Information Retrieval*. ACM Press. 190-197.
- [41] M. D. Jiménez Ruiz "Modelado formal para la representación y evaluación de reglas de asociación," E.T.S. de ingeniería informática y telecomunicación Tesis Doctoral, Departamento de Ciencias de la comunicación e inteligencia artificial Universidad de Granada, Granada, 2010.

- [42] R. Uday Kiran and P. Krishna Re, "An improved multiple minimum support based approach to mine rare association rules," in *Computational Intelligence and Data Mining, 2009. CIDM '09. IEEE Symposium on*, 2009, pp. 340-347.
- [43] M. Kumar and R. Vig, "Focused Crawling Based Upon TF-IDF Semantics and Hub Score Learning," *Journal of Emerging Technologies in Web Intelligence*, vol. 5, pp. 70-77, 2013.
- [44] L. Hyeokju, H. Joon, and K. Sung-Ryul, "Implementation of a Large-Scalable Social Data Analysis System Based on MapReduce," in *Computers, Networks, Systems and Industrial Engineering (CNSI), 2011 First ACIS/JNU International Conference on*, 2011, pp. 228-233.
- [45] W. Wenxian, C. Xingshu, Z. Yongbin, W. Haizhou, and D. Zongkun, "A Focused Crawler Based on Naive Bayes Classifier," in *Intelligent Information Technology and Security Informatics (IITSI), 2010 Third International Symposium on*, 2010, pp. 517-521.
- [46] M. Kudelka, V. Snasel, Z. Horak, and A. E. Hassanien, "Web Communities Defined by Web Page Content," in *Web Intelligence and Intelligent Agent Technology, 2008. WI-IAT '08. IEEE/WIC/ACM International Conference on*, 2008, pp. 385-389.
- [47] A. Khemphila and V. Boonjing, "Comparing performances of logistic regression, decision trees, and neural networks for classifying heart disease patients," in *Computer Information Systems and Industrial Management Applications (CISIM), 2010 International Conference on*, 2010, pp. 193-198.
- [48] Z. Lijuan, W. Linshuang, G. Xuebin, and S. Qian, "A clustering-Based KNN improved algorithm CLKNN for text classification," in *Informatics in Control, Automation and Robotics (CAR), 2010 2nd International Asia Conference on*, 2010, pp. 212-215.
- [49] J. M. Exposto, J. , A. Pina, A. Alves, and J. Rufino. (2005, Geographical Partition for Distributed Web Crawling. *GIR '05: Proc. of the Geographic Information Retrieval*, 55–60.
- [50] T. Hasegawa, S. Sekine, and R. Grishman, "Discovering relations among named entities from large corpora," in *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, 2004, p. 415.
- [51] T. Hirano, Y. Matsuo, and G. Kikui, "Detecting semantic relations between named entities in text using contextual features," in *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, 2007, pp. 157-160.
- [52] M. Lan, "Which Who are They? People Attribute Extraction and Disambiguation in Web Search Results," *ACM09*, 2009.
- [53] J. O. Gong, Douglas, "Determine the Entity Number in Hierarchical Clustering for Web Personal Name Disambiguation," *WWW2009*, 2009.
- [54] J. Artilles, J. Gonzalo, and S. Sekine, "The SemEval-2007 WePS evaluation: establishing a benchmark for the web people search task," presented at the Proceedings of the 4th International Workshop on Semantic Evaluations, Prague, Czech Republic, 2007.
- [55] E. Elmacioglu, Y. F. Tan, S. Yan, M.-Y. Kan, and D. Lee, "PSNUS: web people name disambiguation by simple clustering with rich features," presented at the Proceedings of the 4th International Workshop on Semantic Evaluations, Prague, Czech Republic, 2007.
- [56] N. S.-T. Cristianini, John, *An introduction to support Vector Machines: and other kernel-based learning methods*: Cambridge University Press, 2000.
- [57] B. Furht, *Handbook of Social Network Technologies and Applications*: Springer-Verlag New York, Inc, 2010.



- [58] C. Tobar, A. Germer, J. Adán-Coello, and R. de Freitas, "Retrieving Wiki Content Using an Ontology Mining and Analyzing Social Networks." vol. 288, I. H. Ting, H.-J. Wu, and T.-H. Ho, Eds., ed: Springer Berlin / Heidelberg, 2010, pp. 21-33.
- [59] M. Forestier, J. Velcin, and D. Zighed, "Extracting Social Networks to Understand Interaction," in *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on*, 2011, pp. 213-219.
- [60] P. Bogdanov, N. D. Larusso, and A. Singh, "Towards Community Discovery in Signed Collaborative Interaction Networks," in *Data Mining Workshops (ICDMW), 2010 IEEE International Conference on*, 2010, pp. 288-295.
- [61] J. Roperó, "Método general de Extracción de información basado en el uso de Lógica Borrosa. Aplicación en portales web," Tesis doctoral, Ingeniería Informática, Universidad de Sevilla, España, 2009.
- [62] A. Manchego, "Sistema de información de detección de plagio en documentos digitales usando el método Document Fingerprinting," 2009.
- [63] K. Jung-Tae, L. Jong-Hoon, L. Hoon-Ki, and P. Eui-Hyun, "Design and Implementation of the Location-Based Personalized Social Media Service," in *Internet and Web Applications and Services (ICIW), 2010 Fifth International Conference on*, 2010, pp. 116-121.
- [64] J. Moody and D. R. White, "Structural cohesion and embeddedness: A hierarchical concept of social groups," *American Sociological Review*, vol. 68, pp. 103-127, 2003.
- [65] U. Bandes, "A Faster Algorithm for Betweenness Centrality," *Journal of Mathematical Sociology*, vol. 25, pp. 163-177, 2001.
- [66] M. E. J. Newman, "Mathematics of networks," *The New Palgrave Encyclopedia of Economics*, vol. 2nd edition, 2007.
- [67] A. Bavelas, *A mathematical model for group structures* vol. 7: Massachusetts Institute of Technology, 1948.
- [68] I. Pohl, *An algorithm for finding bridges and its extensions*: Stanford Linear Accelerator Center. Computation Group, 1968.
- [69] D. J. W. y. S. Strogatz, *Collective dynamics of 'small-world' networks*, 1998.
- [70] D. Junhua, I. Cruz, and L. ChengCheng, "A formal model for building a social network," in *Service Operations, Logistics, and Informatics (SOLI), 2011 IEEE International Conference on*, 2011, pp. 237-242.
- [71] K. M. M. Risvik, R, "Search Engines and Web Dynamics," *Computer Networks*, vol. 39, pp. 289-302, 2002.
- [72] M. Shoaib and A. Basharat, "Ontology based knowledge representation and semantic profiling in personalized semantic social networking framework," in *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on*, 2010, pp. 95-99.
- [73] M. Lobur, A. Romanyuk, and M. Romanyshyn, "Using NLTK for educational and scientific purposes," in *CAD Systems in Microelectronics (CADSM), 2011 11th International Conference The Experience of Designing and Application of*, 2011, pp. 426-428.
- [74] I. Amón, "Guía metodológica para la selección de técnicas de depuración de datos," Tesis de Maestría, Facultad de Minas, Escuela de Sistemas, Universidad Nacional de Colombia, Medellín, 2010.

Mónica Andrea Niño Barón

Ingeniera de Sistemas Especialista en gestión de Sistemas y Tecnologías de la Información, estudiante MSc.. Ciencias de la Información y las Comunicaciones y miembro del grupo de investigación GESDATOS en la Universidad Distrital Francisco José de Caldas en Bogotá.
e-mail: ninoandrea75@hotmail.com

Sonia Ordoñez Salinas

Docente de la Universidad Distrital, Facultad de Ingeniería. Estadística de la Universidad Nacional. Ingeniera de Sistemas de la Universidad Distrital. Especialista en Teleinformática, Universidad Distrital. Magíster en Ingeniería de Sistemas, Universidad Nacional. Doctor en Ingeniería de Sistemas y Computación, Universidad Nacional. Grupo de Investigación Gesdatos U.D. sordonez@udistrital.edu.co

Ubicación y dimensionamiento como parámetros en el diseño de almacenes: revisión del estado de arte

Location and sizing as parameters in the warehouse design: A survey paper

**Andrés Mauricio
Hualpa Zúñiga**

Universidad Cooperativa
Bogotá-Colombia
andres.hualpa@campusucc.edu.co

**Carolina Suárez
Roldán**

Universidad Cooperativa
Sede Bogotá – Colombia
carolina.suarez@campusucc.edu.c



Resumen

Este artículo presenta una revisión literaria de los parámetros de ubicación y dimensionamiento en el diseño de almacenes, en el cual se identificaron modelos, métodos, algoritmos y herramientas que se han propuesto para abordar este problema. Se aplicó una metodología de exploración, recolección, clasificación, evaluación y selección de los trabajos publicados desde los años 1970 hasta 2012. El estudio concluye identificando posibles rutas de investigación en esta área de interés, a la luz de la revisión realizada.

Palabras clave: Diseño de almacén, Ubicación, Dimensionamiento.

Abstract

This article presents a survey of the literature concerning location and design as parameters in the design of stores. The survey includes identification of models, methods, algorithms and tools that have been proposed to address this problem. The methodology applied was based on exploration, collection, classification, evaluation and selection of the papers published in the period 1970-2012. The study concludes by identifying possible research routes in this area of interest according to the survey.

Key words: Warehouse design, Location, Sizing.

1. Introducción

El proceso de ubicación y dimensionamiento de almacenes es un factor estratégico al momento de diseñar sistemas logísticos, debido a que se convierte en un componente físico que genera ventajas com-

Fecha recibido: ene. 14/2013
Fecha modificado: may. 17/2013
Fecha aceptado: may. 17/2013

petitivas de reducción de costos, tiempos de respuesta y servicio al cliente. Este caso específico se aplica para sectores industriales especializados en la producción basada en proyectos, como en el sector de la construcción. En él, el diseño de almacenes tiende a convertirse en una actividad variable y temporal que depende de diferentes condiciones específicas de la obra y que se desarrolla en mayor parte a través de la experiencia, dejando de lado el conocimiento y técnicas que pueden generar resultados con mayor eficiencia. A pesar de los avances realizados en el campo de la Ingeniería Industrial relacionados con el proceso de ubicación y dimensionamiento de instalaciones, se está dejando de lado la oportunidad de generar aplicaciones de innovación y desarrollo que determinen mejores prácticas en el proceso operativo de este tipo de empresas. En este contexto se propone la adaptación de técnicas expresadas mediante algoritmos, que faciliten la interpretación de operaciones secuenciales y que determinen soluciones eficientes al diseñar almacenes temporales. En este artículo se realiza una revisión literaria desde la década de los 70 hasta la década actual, además de identificar herramientas y tendencias en el estado del arte, y finaliza con una discusión sobre posibles rutas de investigaciones en el área de interés.

2. Materiales y métodos

Para la construcción del estado del arte se aplicó un método adaptado del modelo de redacción de reseñas bibliográficas [19], integrado por tres etapas generales: Búsqueda, Selección analítica y Extracción de información para desarrollo de artículos de revisión. A su vez, en cada una de estas etapas se utilizaron variables metodológicas que se describen en la Figura 1.

En la etapa de “Búsqueda”, se definieron dos palabras clave para delimitación del tema: “Ubicación” y “Dimensionamiento” de almacenes. Como estrategia de búsqueda se aplicó el procedimiento ascendente, localizando resultados de estudio a partir de referencias bibliográficas. El porcentaje de consulta de fuentes de información es de un 85% con la base de datos *Science Direct* y un 15% con otras bases de datos bibliográficas, entre ellas *Redalyc* y *Proquest*. Adicionalmente a este último porcentaje se integra la consulta de literatura aplicada al tema de instalaciones de almacenamiento. El objeto de identificar la información principalmente en la base de datos *Science Direct* obedece a la rigurosidad y acceso de artículos que contiene, específicamente en la sección de Ciencias físicas e Ingeniería. Un criterio adicional de búsqueda fue bajo los tópicos respectivos a: *Supply Chain, Journal of Production Economics, Operations Research and Mangement, Computers and Industrial Engineering*, dado que muchos resultados tenían relación con las ciencias de la salud y sistemas inteligentes de información. Los resultados obtenidos generaron un total de 164 artículos preseleccionados para la etapa posterior.

En la etapa de “Selección Analítica”, los artículos obtenidos fueron clasificados bajo el criterio de consistencia con el tema de diseño de instalaciones de almacenaje, específicamente con los parámetros de ubicación y dimensionamiento, siendo neces-

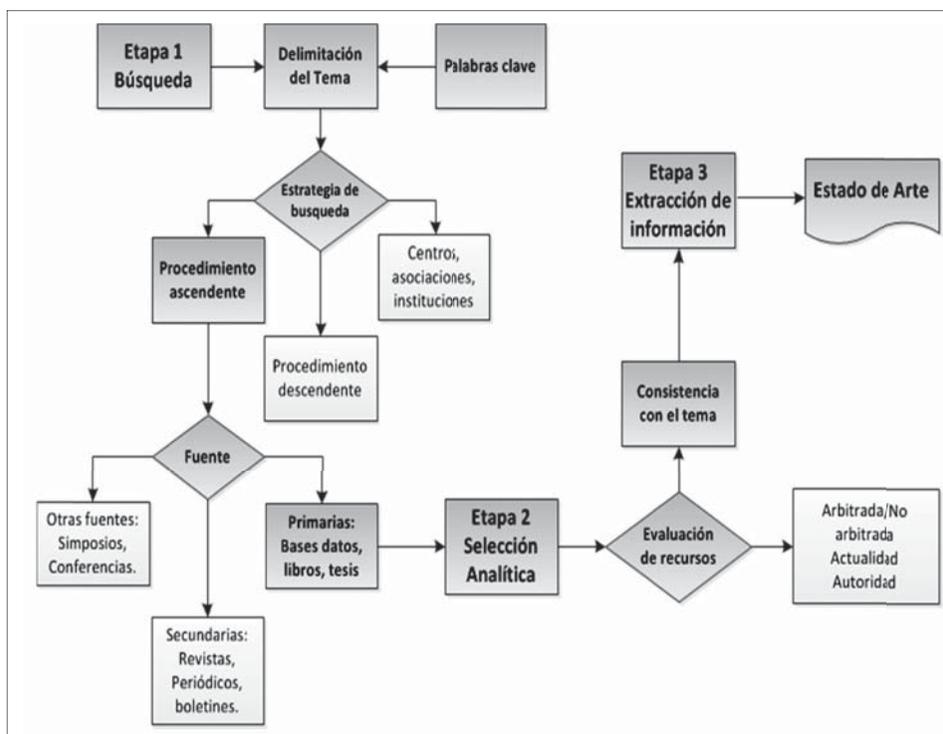


Figura 1. Metodología aplicada para el desarrollo del estado de arte.

ría una exclusión del 64% del total obtenido inicialmente. En la selección de artículos finales se identificó adicionalmente que algunos de ellos tienen en cuenta la combinación del parámetro de distribución.

En la etapa final de “Extracción de información”, los artículos y documentos finales fueron interpretados y organizados en una matriz por año, autor, y método de solución propuesto, siendo este último el principal aspecto de revisión para determinar parámetros de diseño del algoritmo relacionado con el proyecto de investigación de maestría.

3. Estado del arte

El diseño del almacén consta de cinco grandes decisiones: determinación de la estructura, dimensionamiento, diseño interno, selección de equipos y selección de estrategias operacionales [32] [48], representando cada una de ellas una etapa determinante para la integración de una instalación de almacenamiento. Para el caso específico del “diseño interno”, existen dos parámetros claves que son la ubicación de las instalaciones y el dimensionamiento de las mismas, como se expresa en la Figura 2. A pesar que se destaca la in-

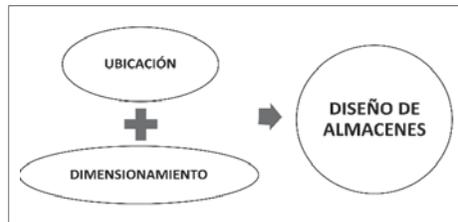


Figura 2. Parámetros para el diseño de un almacén.

Fuente: Autores

terrelación de estos dos aspectos, estudios realizados por diferentes autores a lo largo de 40 décadas generalmente los han tomado por separado. A partir de lo anterior se desarrolla la revisión de literatura obtenida de artículos disponibles en bases de datos y teniendo como referencia de búsqueda el termino: “diseño del almacén”.

Entre los estudios desarrollados se tiene registro desde la década de los 70’s, en donde se encuentran modelos de localización, entre ellos el de múltiples periodos presentando soluciones factibles a partir de programación dinámica [47]. Este modelo presenta restricciones de capacidad acondicionando distintas variables, entre ellas el número máximo de instalaciones por periodo, posibilidad de recolocación, entre otras. Dos años más tarde se plantea la heurística STPL (*Stochastic Transportation Location Problem*) [37], que determina la ubicación de las plantas de acuerdo a la programación de envíos cuyo comportamiento es estocástico. Esta heurística supone que existe un número discreto de posibles ubicaciones y que los costos de envío son lineales.

En la década de los 80’s se plantea un modelo de localización multi-periodo conocido como modelo DUFLP (*Dynámic Uncapacitated Facility Location Problem*) [48]. Este modelo no tiene restricciones de capacidad, permitiendo abrir o cerrar instalaciones. La solución para este modelo es el algoritmo *Dual Ascent Method*, descartando así a la programación dinámica como herramienta imprescindible para dar solución a problemas de localización dinámica. En el año 1984 se acondiciona la formulación de programación lineal para el dimensionamiento óptimo del almacén [41]. En este modelo se extiende para el problema dinámico el tamaño, generando flexibilidad en su diseño. En el año 1986 se plantea un modelo de diseño de almacén basado en los costos de inventario y preparación de pedidos [9].

En la década de los 90’s se desarrolla un estudio orientado al dimensionamiento de la zona de almacenaje en donde se establece que su tamaño depende del costo de almacenamiento y el costo de reposición [36]. Este modelo se formula para el caso de almacenes con estanterías convencionales obteniendo soluciones a partir de aproximación estocástica. En el mismo año se propone una ampliación a este modelo mediante la integración del índice cúbico por pedido (COI), como parámetro de asignación de políticas de inventario y preparación de pedidos [10] [11]. Desde este año se visualiza el incremento en la utilización de aplicaciones computacionales para este proceso. Mediante la combinación de modelos matemáticos y computacionales se desarrollan estudios de simulación en actividades de automatización de almacenes [43]. Entre los principales problemas trabajados en este estudio se contempla la baja utilización de espacios disponibles para almacenamiento. Se propone el incremento de la capacidad del almacén mediante la utilización de espacios ubicados en altura y la asignación de rutas de recolección de pedi-



dos mediante teoría de grafos. Un año después se desarrolla el modelo DCPLP (*Dynamic Capacitated Facility location problema*) [1], el cual retoma restricciones de capacidad finita con la posibilidad de aumentar la misma. La solución propuesta fue desarrollada a partir de técnicas lagrangianas aplicadas a límites superiores e inferiores mediante un método basado en programación dinámica.

En el año 1992 se realiza una revisión de los modelos aplicados al diseño de la capacidad de almacenamiento [23]. En dicho estudio se determina que la capacidad de almacenamiento depende de cuatro factores: políticas de alistamiento de pedido, procesamiento de lote, asignación de almacenamiento y modelos de control dinámico. En el año 1996 se retoma el estudio del comportamiento aleatorio de la demanda como un parámetro de alistamiento de materiales y de diseño de la zona de almacenamiento. Mediante el modelo de evaluación de sistemas integrados de almacenamiento [8], se plantea el análisis de costos asociados con el diseño y operación de un sistema de distribución. Para el año 1998 se desarrolla el algoritmo UFLP (*Uncapacitated Facility Location Problem*) [4], para ubicación de instalaciones con restricciones de capacidad a partir del método *Branch and Bound*. En el mismo año fue desarrollado un modelo de localización discreta mediante la técnica heurística de *DROP* apoyado de la herramienta *CPLEX* [14], considerando todas las instalaciones iguales con capacidad ilimitada. En este mismo año se desarrolla un modelo de programación lineal para el dimensionamiento económico de un almacén en el que las demandas presentan comportamiento estacional [2]. Mediante la formulación de este modelo se obtiene el tamaño óptimo de la bodega. Un año después se presenta una clasificación de los problemas de gestión de almacenes [28]. En este trabajo se realiza una revisión frente al diseño del sistema de gestión de inventario teniendo en cuenta restricciones de espacio. Posteriormente se agrega la novedad de considerar la manipulación de múltiples productos ubicados en almacenes intermedios que cumplen restricciones de producción y demanda [58]. Su función objetivo busca minimizar el costo de la instalación. En este mismo año se desarrolló el modelo denominado CFLP (*Capacitated Facility Location Problem*) [3], el cual se caracteriza por trabajar la combinación de localización y capacidad suponiendo que las instalaciones no tienen una capacidad fija ya que existen variables de decisión que integran dimensiones de cada una de las instalaciones. De la misma forma se obtiene un modelo matemático para dimensionamiento basado en el flujo de materiales, el cual integra la configuración del almacenamiento y los parámetros para alistamiento de pedidos. A partir de ello se presenta como resultado ciclos de tiempo para realizar el procesamiento de pedidos. En este mismo año se presenta una revisión frente al proceso de diseño de almacenes [5], en el que se propone un proceso conformado por cinco fases: Concepto y adquisición de datos, especificaciones funcionales, especificaciones técnicas, selección de los equipos y selección de las políticas de planificación y control.

Dos años más tarde se realizó una modificación al modelo CFLP creando el modelo MCLFP (*Multiple Capacitated Facility Location Problem*) [42], en donde se incluye restricciones de tiempo para un único periodo y su solución está definida mediante técnicas meta heurísticas y de simulación. En el año 2003 se desarrolla un modelo heurístico que deter-

mina el menor número de zonas de almacenamiento necesarios [46]. Este modelo integra múltiples productos y a cada uno un depósito con capacidad limitada. En el año 2005 se presenta un método de dimensionamiento de almacenes basado en políticas de reordenación de materiales [20]. Este se desarrolla en tres etapas de medición. En este mismo año se trabaja el problema combinado de ubicación y asignación óptima de inventarios [39]. La solución propuesta se desarrolla mediante heurística de búsqueda exhaustiva PODA. En el año 2006 se desarrolla un modelo de evaluación de desempeño de la configuración de un almacén [27], el cual está diseñado para mejorar el rendimiento al momento de preparar un pedido. Los parámetros de comparación están determinados por el tamaño del almacén, el tamaño de los estantes y el número de transportadores. En este mismo año se desarrolla un modelo de asignación de materiales solucionado mediante programación dinámica [25]. La aplicación de este modelo fue realizada en una industria de materiales de acero.

En el año 2007 se plantea un modelo de dimensionamiento del almacén a partir del sistema de disposición interna y los métodos de enrutamiento para el procesamiento de pedidos [45]. En el mismo año se plantea un modelo de dimensionamiento del almacén considerando las restricciones de capacidad obtenida a partir de un proceso de planeación de la producción [12]. Este modelo considera la distribución de múltiples productos a lo largo de una cadena de suministro teniendo la posibilidad de activar o desactivar el almacén según se requiera. Su solución está definida a partir de la combinación de métodos de programación lineal, entera mixta y relajación lagrangiana. De la misma manera se realiza una revisión sobre el grado de relación que existe entre las funciones básicas del almacén: recepción, almacenamiento, preparación de pedidos, envíos y el dimensionamiento del mismo [31]. El proceso de ubicación de almacenes vuelve a trabajarse en este año mediante la combinación de las metodologías AHP (*Analytic Hierarchy Process*) y DEA (*Data Envelopment Analysis*) [33]. El resultado determina que el orden de aplicación de las metodologías es generar rendimientos de cada ubicación del almacén mediante AHP, como información de entrada para medir la eficiencia de los mismos mediante DEA. En este mismo año se desarrolla un modelo de distribución de zonas de almacenamiento a partir del índice cúbico de ocupación [56]. Esto determina la clasificación de materiales para la asignación de espacios de almacenamiento.

En el año siguiente se desarrolla una configuración de almacenes de varios niveles que minimiza el costo de mantenimiento anual [51]. Este algoritmo PSO (*Particle Swarm Optimization*) integra problemas de distribución, orientación de bastidores de almacenamiento, asignación de espacios, que ha demostrado ser un modelo matemático NP-duro. En este mismo año se aplica el algoritmo *Branch and Bound* en el dimensionamiento de un almacén basado en los costos de manejo de materiales [55]. Sus resultados demostraron ser más eficientes que las soluciones mediante programación dinámica o lineal. En este mismo año se propone un modelo de dimensionamiento del almacén a partir del proceso de reposición de materiales [30]. Los resultados obtenidos a partir del modelo *Pick to belt* demuestran la reducción de esperas en el proceso de abastecimiento.



En el año 2009 se plantea el diseño del almacén a partir de la comparación de tres configuraciones evaluadas con indicadores de distancia de viaje [38]. Los resultados obtenidos demuestran que la designación de un pasillo transversal medio, facilita el proceso de alistamiento y determina las dimensiones mínimas del almacén. En este mismo año se plantea la integración de un modelo sistemático para el diseño de almacenes [44]. En este modelo se tiene en cuenta la integración de herramientas y técnicas resumidas en 11 pasos. De la misma manera la integración del modelo EOQ como parámetro de dimensionamiento [7]. En tal estudio la capacidad de almacenamiento máximo es una variable de decisión. Por otra parte, en ese año se realiza la comparación de tres heurísticas que determinan la temporización de reposición de artículos teniendo en cuenta las restricciones de capacidad en el almacén [52]. Los resultados obtenidos en esta comparación demostraron que aplicando una política de reposición de lote, por lote en horarios definidos, se generan ahorros en los costos asociados. Este método de reposición a su vez define la configuración del almacén. Por otra parte el modelo de distribución física “espina de pescado” también ha sido evaluado [53]. Allí se realiza la comparación con el modelo tradicional de cuadrícula en donde se obtiene mayor rendimiento en términos de distancia de recorrido y utilización de espacio.

En el año 2010 se desarrolla un estudio en donde se analiza la relación cuantitativa que existe entre los límites de capacidad para el dimensionamiento del almacén y las decisiones de inventario [22]. Se identifica a través de este estudio que los costos de inventario se pueden reducir teniendo en cuenta la planificación y restricciones de almacenamiento. En este mismo año se elabora una revisión de los métodos cuantitativos al diseño de almacenes [24]. En sus resultados principales se propone la utilización de técnicas de simulación y diseño experimental. Por otra parte se determina la integración del proceso de planeación con el dimensionamiento del almacén [21]

En el año 2011 se desarrolla una investigación que establece que el diseño del almacén estará asociado a variables de dimensionamiento, sistema de almacenamiento, preparación de pedidos y tendencias de la demanda las cuales fueron previamente trabajadas por otros autores [17]. La principal diferencia es que integra otras variables adicionales relacionadas con las características dimensionales del material y tasas de rotación. Los resultados obtenidos a partir de simulación determinaron que el rendimiento de las operaciones del almacén depende totalmente de estas variables combinadas con la asignación multinivel de estanterías. Para este mismo año se adelantó un estudio que evalúa la relación que existe entre el diseño y configuración del almacén y el proceso de alistamiento de pedidos [35]. Se realizó experimentación con diferentes configuraciones de dimensionamiento y el modelo LSO (*Linear Sequence Optimization*), obteniendo como resultado una serie de patrones de secuencia asignados para diferentes dimensiones de almacenamiento. En este mismo año se desarrolló un modelo de dimensionamiento solucionado mediante Algoritmo de búsqueda Tabú [40]. De la misma manera se plantea el método MCDA (*Multiple Criteria Decision Making*) [29], el cual está sujeto a un análisis de decisión multicriterios para la evaluación de la utilidad de un almacén.

Finalmente para el año 2012 se integran varios estudios y aplicaciones de diseño de almacenes. Uno de ellos determina que la configuración “Flying-V” [34], genera resultados numéricos favorables para rendimientos de tiempo y distancia. Esta configuración establece de la misma manera parámetros de dimensionamiento del almacén. En este mismo año se retoma un estudio de localización de almacenes a través de programación estocástica [16]. Este modelo plantea el rediseño de una red de almacenes realizando eliminación de instalaciones. La solución se plantea mediante la aproximación SAA (*Sample Average Approximation*). Los resultados demuestran que las soluciones estocásticas tienen mayor solidez que las soluciones deterministas. Otro parámetro propuesto para el diseño de almacenes es la inclusión del proceso de asignación de materiales en instalaciones disponibles [15]. Este estudio plantea la comparación de las soluciones obtenidas entre un modelo matemático lineal y una heurística de búsqueda local iterativa. El objetivo planteado es reducir tiempos de entrega, niveles de inventario, costos de almacenamiento e incrementar el nivel de servicio. El resultado demuestra que la heurística genera buenos resultados en menor tiempo y esfuerzo computacional de cálculo. Nuevamente se plantea un modelo de dimensionamiento y configuración del almacén basado en el comportamiento estocástico de la demanda [59]. En este trabajo se determina como factor fundamental de diseño el grado de relación entre el sistema de distribución de materiales y el dimensionamiento del almacén. Los casos que se analizan especialmente son los de flujo de material de almacenamiento con retorno. La solución se plantea en cuatro partes a la que se aplica modelos de programación lineal. También se plantea un modelo de dimensionamiento del almacén a partir de la configuración de la instalación en “U” [13]. De acuerdo con este estudio se determina que este tipo de distribución minimiza la distancia total de recorrido y por ende determina algunas condiciones de longitud del almacén. El modelo de asignación de almacenamiento por carga unitaria [26] [6] se convierte en otro modelo de dimensionamiento dependiente de la configuración de la instalación. Para este caso el indicador a ser evaluado se relaciona con el tiempo de ocupación para movimiento de materiales y tiempos de respuesta. De la misma manera otro estudio que evalúa el indicador de tiempo de respuesta es el algoritmo *On-line Order Picking* [49] [50]. Este algoritmo evalúa la aplicación de diferentes heurísticas de ruteo combinadas con la asignación de materiales.

4. Resultados

Los artículos obtenidos para la construcción del estado de arte fueron ordenados de acuerdo al año de publicación y su autor respectivo. A partir de esto se identificaron tres parámetros particulares de análisis: ubicación, dimensionamiento y distribución. En la Tabla I se integran los resultados para la interpretación y construcción del algoritmo de asignación de espacios de almacenamiento teniendo en cuenta para cada uno de los parámetros definidos los resultados y el método de solución propuesto por cada autor.

**Tabla I.** Evolución de la literatura en el proceso de diseño de almacenes

Año	Autor	Resultado	Parámetro de diseño	Método de solución
1975	Ballou [47]	Modelo Localización Dinámica	Ubicación	Programación dinámica
1977	Larry [37]	Heurística STPL	Ubicación	Heurística STPL
1982	Roy y Erlenkotter [48]	Modelo DUFLP	Ubicación	Algoritmo DAM
1984	Hung y Fisk [41]	Modelo de dimensionamiento de almacén basado en modelo de programación lineal	Dimensionamiento	Método Simplex
1986	Malmborg, Balachandran y Kyle [9]	Modelo de diseño de almacén basado en costos de inventario & Picking	Dimensionamiento	Procedimiento Heurístico
1990	Bhaskaran y Malmborg [36]	Modelo de diseño de almacén basado en costos de almacenamiento y reposición.	Dimensionamiento	Aproximación estocástica
1990	Malmborg y Bhaskaran [10]	Modelo de diseño de almacén basado en índice cúbico por pedido.	Dimensionamiento	Aplicación computacional
1990	Hounkonnou [43]	Modelo de diseño de almacén basado en la utilización de la capacidad.	Dimensionamiento	Simulación
1991	Shulman [1]	Modelo DCPLP	Ubicación	Relajación Lagrangiana
1992	Cormier y Gunn [23]	Revisión modelos de capacidad de almacenamiento	Dimensionamiento	Programación lineal
1996	Malmborg [8]	Modelo de evaluación de sistemas integrados de almacenamiento	Dimensionamiento	Aplicación computacional
1998	Klose [4]	Algoritmo UFLP	Ubicación	Método Branch and Bound
1998	Saldanha y Captivo [14]	Modelo de localización discreta	Ubicación	Heurística DROP
1998	Rao [2]	Modelo de dimensionamiento de almacén basado en modelo de programación lineal	Dimensionamiento	Método Simplex
1999	Van den Berg y Zijm [28]	Diseño de sistema de gestión de inventario sujeto a restricciones de espacio	Dimensionamiento y Distribución.	Aplicación computacional
2000	Hinojosa y Puerto [58]	Modelo de localización multiproducto	Ubicación	Programación Lineal

Año	Autor	Resultado	Parámetro de diseño	Método de solución
2000	Klose y Drexl [3]	Algoritmo CFLP	Ubicación y Dimensionamiento	Modelación Matemática
2000	Rouwenhorst, Reuter, Stocakrahm, Houtum, Mantel, y Zijm [5]	Revisión diseño de almacén	Dimensionamiento y Distribución.	5 fases de diseño
2002	Ortega y Delgado [42]	Modelo MCLFP	Ubicación y Dimensionamiento	Simulación
2003	Kalfakakou, Katsavounis y Tsouros [46]	Modelo de ubicación de almacén	Ubicación	Teoría de grafos
2005	Broulias, Marcoulaki, Chondrocoukis y Laios [20]	Modelo de dimensionamiento basado en políticas de reordenamiento de materiales	Dimensionamiento	Método de tres etapas
2005	Moreno, Velasquez y Díaz [39]	Modelo de Ubicación y asignación de Inventarios	Ubicación y Dimensionamiento	Búsqueda exhaustiva
2006	Hwang y Cho [27]	Modelo de dimensionamiento de almacén	Dimensionamiento	Simulación gráfica
2006	Günther Zäpfel y Wasner [25]	Modelo de dimensionamiento a través de reposición de materiales	Dimensionamiento	Programación dinámica
2007	Koster, Le-Duc y Roodbergen [45]	Modelo de diseño de almacén basado en la disposición interna y métodos de enrutamiento.	Dimensionamiento	Modelación Matemática y simulación
2007	Aghezzaf [12]	Modelo de diseño de almacén basado en la distribución de múltiples productos.	Ubicación	Programación Lineal, entera mixta y relajación lagrangiana
2007	Gu, Goetschalckx y McGinnis [31]	Modelo de diseño de almacén basado en funciones básicas del almacén.	Dimensionamiento	Programación Lineal
2007	Korpela, Lehmusvaara y Nisonen [33]	Modelo de Ubicación de almacén mediante combinación de metodologías AHP y DEA	Ubicación	Proceso analítico jerárquico AHP y análisis envolvente de datos DEA



Año	Autor	Resultado	Parámetro de diseño	Método de solución
2007	Muppani y Adil [56]	Modelo de distribución de materiales	Dimensionamiento y Distribución.	Simulación
2008	Önüt, Tuzkaya y Doğaç [51]	Algoritmo Dimensionamiento de almacén	Dimensionamiento	Modelación matemática NP-duro
2008	Muppani y Adil [55]	Algoritmo Dimensionamiento de almacén	Dimensionamiento	Algoritmo Branch and Bound
2008	Gagliardi, Ruiz y Renaud [30]	Modelo de dimensionamiento a través de reposición de materiales	Dimensionamiento y Distribución.	Algoritmo Pick to belt
2009	Pohl, Meller y Gue [38]	Modelo de diseño de almacén basado en indicadores de distancia.	Dimensionamiento y Distribución.	Modelación Matemática
2009	Baker y Canessa [44]	Modelo de diseño de almacén basado en configuración interna.	Dimensionamiento y Distribución.	Algoritmo en cascada
2009	Minner [52]	Modelo de Diseño de almacén basado en Dinámica de la demanda con capacidad de almacenamiento limitada.	Dimensionamiento	Comparison of Constructive, savings and Smoothing heuristics
2009	Cheng, Kotov y Kovalyov [7]	Modelo de dimensionamiento de almacén basado en el modelo EOQ	Dimensionamiento	EOQ
2009	Steffey [53]	Modelo de dimensionamiento y distribución	Dimensionamiento y Distribución.	Evaluación distribución espina de pescado
2010	Strack y Pochet [22]	Modelo de diseño de almacén basado en Relación cuantitativa de Inventario.	Dimensionamiento	Programación Lineal
2010	Gómez y Correa [24]	Revisión métodos cuantitativos para el diseño de almacenes	Dimensionamiento	Simulación y métodos cuantitativos
2010	Geraldes, Sameiro y Pereira [21]	Modelo integral de gestión de almacén	Dimensionamiento y Distribución.	Integración de Modelos de Decisión
2011	F. Chan y H.K. Chan [17]	Modelo de diseño de almacén basado en variables de demanda, Picking y sistema de almacenamiento.	Dimensionamiento	Simulación
2011	Moeller [35]	Modelo de diseño de almacén basado en variables Picking.	Dimensionamiento	Método LSO
2011	Chen, Langevin y Riopel [40]	Modelo de dimensionamiento basado en la reubicación de materiales	Dimensionamiento	Búsqueda Tabú

Año	Autor	Resultado	Parámetro de diseño	Método de solución
2011	ZAK y SAWICKI [29]	Método MCDA para diseño de almacenes	Dimensionamiento	Método MCDA
2012	Gue, Ivanović y Meller [34]	Modelo de diseño de almacén basado en configuración Fying	Dimensionamiento y Distribución.	Modelación Matemática y simulación
2012	Kiya y Davoudpour [16]	Modelo Localización Estocástica	Ubicación	Aproximación SAA y Descomposición de Benders
2012	Guerriero, Musmanno, Pisacane y Rende [15]	Modelo de diseño de almacén basado asignación de materiales e instalaciones disponibles.	Dimensionamiento	Heurística de Búsqueda local iterativa
2012	Chou, Y. Chen y H. Chen [59]	Modelo de diseño de almacén basado en demanda estocástica	Dimensionamiento	Programación estocástica
2012	Cakmak, Gunay, Aybakan y Tanyas [13]	Modelo de diseño de almacén basado en configuración U	Dimensionamiento y Distribución.	Modelación Matemática y simulación
2012	Carlo y Giraldo [26]	Modelo de diseño de almacén basado en configuración de carga unitaria	Dimensionamiento y Distribución.	Modelación Matemática y simulación
2012	Henn [50]	Modelo de diseño de almacén basado en indicadores de tiempo.	Dimensionamiento y Distribución.	Algorithm On-Line Order Picking

A partir de la consulta realizada se puede identificar que en el desarrollo investigativo que ha tenido el proceso de Diseño de Almacenes a lo largo de cuatro décadas, tiene como punto de partida la aplicación de modelos y técnicas matemáticas de localización de instalaciones inicialmente. Este comportamiento exclusivo se mantiene sólo hasta finales de la década de los 80, periodo en el cual comienza a trabajarse el dimensionamiento mediante algoritmos de programación lineal y programación estocástica, orientados a la minimización de costos de inventario y optimización del proceso de alistamiento de pedidos. Desde la década de los 80 en adelante el dimensionamiento, la ubicación y distribución han sido parámetros de diseño que se han trabajado hasta la fecha, caracterizados especialmente por la utilización de aplicaciones computacionales fundamentadas en modelos matemáticos de programación lineal, estocástica y dinámica, algoritmos de búsqueda y simulación. En la Figura 3 se presenta la ventana de tiempo en la que se caracteriza la evolución de publicaciones relacionadas con el diseño de almacenes, teniendo como parámetros de clasificación la ubicación, dimensionamiento y distribución, en donde en algunos casos se trabajaron de una manera combinada.

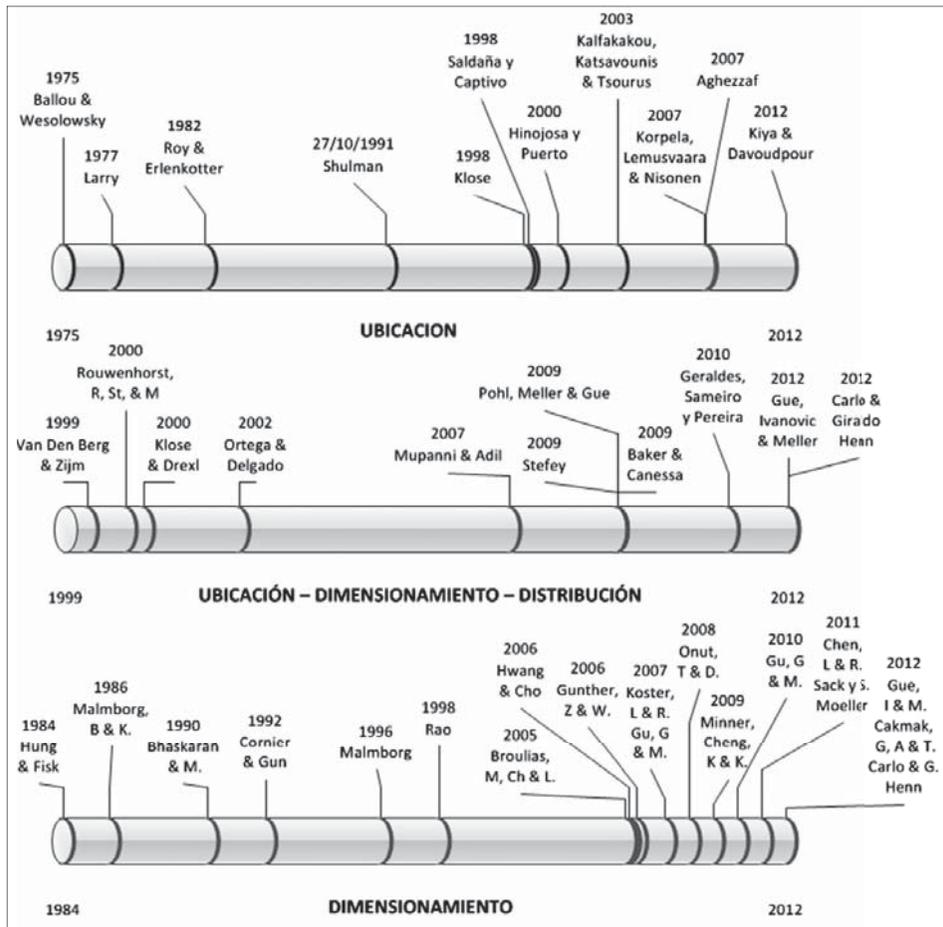


Figura 3. Parámetros para el diseño de un almacén.

5. Discusión de resultados

Luego de la revisión del estado de arte sobre el diseño de instalaciones de almacenamiento, se identifica que ubicación y dimensionamiento no son los únicos parámetros que han sido objeto de investigación y desarrollo. Varios expertos agregan el parámetro de distribución al proceso de gestión del almacén, dado que se combina con el dimensionamiento. Según la Tabla II, los resultados demuestran que el 49%

Parámetro	%
Dimensionamiento	49%
Dimensionamiento y Distribución	24%
Ubicación	21%
Ubicación y Dimensionamiento	6%

de los estudios se enmarcan en su mayoría en el desarrollo de métodos matemáticos, heurísticos y algorítmicos para el dimensionamiento del almacén, seguido por un 24% por la combinación entre este y el parámetro de distribución. De la misma manera el parámetro de ubicación ocupa el tercer lugar con un 21% de participación. Adicionalmente la combinación entre los parámetros de ubicación y dimensionamiento ocupan el nivel más bajo de la tabla, con un 6% de participación.

De acuerdo con la revisión se identifica un bajo desarrollo investigativo en la integración total de los parámetros de diseño de almacenes. Generalmente la ubicación, dimensionamiento y distribución se toman por separado ocasionando incertidumbre en cuanto a la objetividad y globalidad de los resultados. Una de las principales causas está asociada a la parcialidad de los investigadores debido a la complejidad del funcionamiento de los almacenes, representado en la planeación y ubicación de espacios de los procesos de recepción, manejo y custodia de materiales, alistamiento de pedidos, acondicionamiento de unidades de carga, expedición entre otros, así como la relación de variables de tipo de producto, el comportamiento de la demanda, el nivel de servicio al cliente, el flujo de materiales y la flexibilidad del almacén. A lo anterior se agrega una nueva característica que están asumiendo los almacenes, donde las unidades no son enviadas a la zona de almacenamiento sino acondicionadas en un área de expedición para ser distribuidas de la manera más inmediata.

Los modelos de dimensionamiento del almacén resuelven el problema de la capacidad bajo un criterio de costos de construcción y operación, omitiendo factores de importancia como es el manejo de materiales; es decir que es necesario desarrollar modelos más robustos que junto con el desarrollo de hardware, software de optimización y simulación permitan considerar los efectos del dimensionamiento frente a variables y/o parámetros definidos en el diseño integral del almacén y que respondan a problemas que no han recibido ninguna atención, ya que en la revisión no se encontraron estudios que consideren tales parámetros.

El diseño del almacén ha sido un tema de investigación abordado por varios autores, con diferentes propuestas que varían el alcance y las actividades que se definen en cada una de las etapas de diseño. A partir de esto el problema sigue siendo un sistema complejo, de manera que buscar una solución óptima frente a la cantidad de posibilidades que se encuentran en cada uno de las etapas del diseño del almacén, es una oportunidad de investigación. Para terminar es importante recalcar que las decisiones que se tomen en el diseño del almacén no pueden estar soportadas solamente en la intuición y experiencia, puesto que los costos de operación y almacenamiento se determinan en gran medida en la fase del diseño y desempeñan un papel vital en el funcionamiento de las cadenas de suministro. De allí la importancia de desarrollar modelos formales que incluyan en su planteamiento tanto reglas de optimización matemática como heurísticas que generen soluciones integrales.



Finalmente se identificó en la revisión que pocos han sido los estudios especializados en el diseño de almacenes en Colombia, pues los resultados se orientan a resolver problemas globales de la cadena de abastecimiento, en las que se integran proveedores, planta de producción, centros de almacenamiento, distribución y clientes; por lo anterior, se considera esta una área de interés para futuras investigaciones en el país.

6. Conclusiones

Particularmente se identifica que el parámetro de dimensionamiento y distribución son predominantes en relación con la problemática prevista para el proceso de asignación de espacios de almacenamiento, debido a que las variables que se trabajan en los diferentes artículos son relacionadas con modelos geométricos y se desarrollan a través de aplicaciones computacionales. Mediante el análisis de cada uno de los trabajos investigativos de dimensionamiento, se resaltan los resultados obtenidos al aplicar políticas de ubicación de materiales de acuerdo al índice cúbico por pedido. Para este caso la metodología propuesta establece la necesidad de tener una estandarización de las dimensiones de los pedidos con el objeto de determinar la capacidad óptima a lo largo de un periodo de planeación. De la misma manera esta capacidad se convierte en un factor dimensional que dependerá del tamaño de los materiales a almacenar, siendo estos necesariamente compatibles con los diferentes equipos e instalaciones con que cuenta el almacén. Cada uno de estos materiales estará sujeto a una dinámica de flujo en función de la demanda, que para algunos casos tendrán componentes probabilísticos. Esta afirmación se obtiene de aquellos modelos de dimensionamiento en donde los métodos de solución propuestos están apoyados de programación estocástica y lineal. Las aplicaciones computacionales son relevantes, debido a la utilización de herramientas informáticas de simulación de datos y gráfica, como es el caso del software *Automod* y *CAD*, en el cual demostraron ser elementos potencialmente eficientes para determinar dimensiones necesarias de almacenamiento, facilitando la medición de rendimientos en distancia de viaje y tiempo de alistamiento de pedidos. Por otra parte se identifica que la configuración interna del almacén es un factor condicional para los casos asociados a restricciones de espacio y/o capacidad. De acuerdo con el análisis teórico, el modelo de reposición de materiales que se utilice se convierte en una solución a esta problemática, ya que puede lograr una compensación en la demanda.

En síntesis se puede confirmar que hay muchos factores que afectan la asignación de espacios de almacenamiento, como es el tamaño, el diseño del sistema de almacenamiento, las características del producto, las tendencias de la demanda, las tasas de rotación, entre otros. Aún más el parámetro de ubicación, a pesar de estar claramente definido entre las primeras etapas de diseño de almacenes, se tiende a descartar su teoría en razón a que no existe una relación directa con la problemática de asignación de espacios de almacenamiento. Esto debido a que los modelos desarrollados se orientan a una macro-localización, siendo esta una aplicación para el diseño de sistemas de abastecimiento y distribución.

7. Referencias

- [1]. A. Shulman. (1991) Algorithm for solving dynamic capacitated plant location problems with discrete expansion sizes. *Operations research*. vol. 39 (3) 423-436.
- [2]. A.K. Rao & M.R. Rao. (1998). Solution procedures for sizing of warehouses. *European Journal of Operational Research*, Vol 108 (1), 16-25.
- [3]. Andreas, Klose y A. Drexl. (2000) Lower bounds for the capacitated facility location problem based on column generation, *Universidad St. Gallen*, 1-3.
- [4]. Andreas Klose, (1998) A branch and bound algorithm for an uncapacitated facility location problem with a side constraint, *International Transactions in Operational Research*, Vol 5, 155-168
- [5]. B. Rouwenhorst, B. Reuter, V. Stockrahm, G.J. van Houtum, R.J. Mantel, W.H.M. Zijm, (2000) Warehouse design and control: Framework and literature review, *European Journal of Operational Research*, Vol 122, Pág. 515-533.
- [6]. C.H.Y. Lam, K.L. Choy, G.T.S. Ho, S.H. Chung, (2012) A hybrid case-GA-based decision support model for warehouse operation in fulfilling cross-border orders, *Expert Systems with Applications*, Vol 39, Pág 7015-7028.
- [7]. C.T. Ng, T.C.E. Cheng, Vladimir Kotov, Mikhail Y. Kovalyov, (2009) The EOQ problem with decidable warehouse capacity: Analysis, solution approaches and applications, *Discrete Applied Mathematics*, Vol 157, Pág 1806-1824.
- [8]. Charles J. Malmborg, (1996) An integrated storage system evaluation model, *Applied Mathematical Modelling*, Vol 20, Pág 359-370.
- [9]. Charles J. Malmborg, Swaminathan Balachandran, Daniel M. Kyle, (1986) A model based evaluation of a commonly used rule of thumb for warehouse layout, *Applied Mathematical Modelling*, Vol 10, Pág 133-138.
- [10]. Charles J. Malmborg, Krishnakumar Bhaskaran, (1990). A revised proof of optimality for the cube-per-order index rule for stored item location, *Applied Mathematical Modelling*, Vol 14, Pág 87-95.
- [11]. Charles J. Malmborg, Khalid Al-Tassan, (2000). An integrated performance model for order-picking systems with randomized storage, *Applied Mathematical Modelling*, Vol 24, Pág 95-111.
- [12]. El-Houssaine Aghezzaf, (2007) Production planning and warehouse management in supply networks with inter-facility mold transfers, *European Journal of Operational Research*, Vol 182, Pág 1122-1139.
- [13]. Emre Cakmak, Noyan Sebla Gunay, Gulsah Aybakan, Mehmet Tanyas, (2012) Determining the Size and Design of Flow Type and U-Type Warehouses, *Procedia - Social and Behavioral Sciences*, Vol 58, Pág 1425-1433
- [14]. F. Saldanha y F., Captivo, (1998) Heuristic approach for the discrete dynamic location problem. *Location science*, vol. 6, 211-223
- [15]. F. Guerriero, R. Musmanno, O. Pisacane, F. Rende, (2012) A mathematical model for the Multi-Levels Product Allocation Problem in a warehouse with compatibility constraints, *Applied Mathematical Modelling*.
- [16]. Farhad Kiya, Hamid Davoudpour, (2012) Stochastic programming approach to re-designing a warehouse network under uncertainty, *Transportation Research Part E: Logistics and Transportation Review*, Vol 48, Pág. 919-936.



- [17]. Felix T.S. Chan, H.K. Chan, (2011) Improving the productivity of order picking of a manual-pick and multi-level rack distribution warehouse through the implementation of class-based storage, *Expert Systems with Applications*, Vol 38 Pág. 2686-2700.
- [18]. Francesco Basile, Pasquale Chiacchio, Domenico Del Grosso, (2012) A control oriented model for manual-pick warehouses, *Control Engineering Practice*, Vol 20, Pág. 1426-1437.
- [19]. Galvan, Jose L. (2012). *Writing Literature Reviews: A guide for students of the social and behavioral Sciences*. Pycszak Publishing. 5 Ed. 81-96
- [20]. G.P. Broulias, E.C. Marcoulaki, G.P. Chondrocoukis and L.G. Laios. Warehouse management for improved order picking performance: an application case study from the wood industry
- [21]. Geraldes, Carla, Sameiro, Maria y Pereira Guilherme. (2010) Integrated models for warehouse design and planing. *Semana de Enganarhia 2010*. Guimaraes
- [22]. Géraldine Strack, Yves Pochet, (2010) .An integrated model for warehouse and inventory planning, *European Journal of Operational Research*, Vol 204, Pág. 35-50.
- [23]. Gilles Cormier, Eldon A. Gunn. (1992). A review of warehouse models, *European Journal of Operational Research*, Vol 58, Pág. 3-13
- [24]. Gómez M., Rodrigo A. Correa E., Alexander A. (2010). Métodos cuantitativos utilizados en el diseño de almacenes y centros de distribución. *Revista Avances en Sistemas e Informática*, Diciembre Pág. 109-117.
- [25]. Günther Zäpfel, Michael Wasner, (2006) Warehouse sequencing in the steel supply chain as a generalized job shop model, *International Journal of Production Economics*, Vol 104, Pág. 482-501
- [26]. Héctor J. Carlo, Germán E. Giraldo, (2012). Toward perpetually organized unit-load warehouses, *Computers & Industrial Engineering*, Vol 63, Pág. 1003-1012.
- [27]. Heung Suk Hwang, Gyu Sung Cho, (2006) A performance evaluation model for order picking warehouse design, *Computers & Industrial Engineering*, Vol 51, Pág 335-342.
- [28]. J.P.van den Berg, W.H.M. Zijm, (1999) Models for warehouse management: Classification and examples, *International Journal of Production Economics*, Vol 59, Pág. 519-528.
- [29]. Jacek Łak a, Piotr Sawicki. (2011). The multiobjective ranking of the warehouse infrastructure in the physical distribution system.
- [30]. Jean-Philippe Gagliardi, Angel Ruiz, Jacques Renaud. (2008) Space allocation and stock replenishment synchronization in a distribution center, *International Journal of Production Economics*, Vol 115, Pág. 19-27.
- [31]. Jinxiang Gu, Marc Goetschalckx, Leon F. McGinnis, (2007) Research on warehouse operation: A comprehensive review, *European Journal of Operational Research*, Vol 177, Pág. 1-21.
- [32]. Jinxiang Gu, Marc Goetschalckx, Leon F. McGinnis, (2010) Research on warehouse design and performance evaluation: A comprehensive review, *European Journal of Operational Research*, Vol 203, Pág. 539-549.
- [33]. Jukka Korpela, Antti Lehmusvaara, Jukka Nisonen. (2007) Warehouse operator selection by combining AHP and DEA methodologies, *International Journal of Production Economics*, Vol 108, Pág. 135-142.
- [34]. Kevin R. Gue, Goran Ivanović, Russell D. Meller. (2012) A unit-load warehouse with multiple pickup and deposit points and non-traditional aisles, *Transportation Research Part E: Logistics and Transportation Review*, Vol 48, Pág. 795-806.
- [35]. Klaus Moeller, (2011) Increasing warehouse order picking performance by sequence optimization, *Procedia - Social and Behavioral Sciences*, Vol 20, Pág. 177-185.

- [36]. Kumar Bhaskaran, Charles J. Malmborg, (1990) Economic tradeoffs in sizing warehouse reserve storage area, *Applied Mathematical Modelling*, Vol 14, Pág 381-385.
- [37]. Larry J. LeBlanc, (1977). A heuristic approach for large scale discrete stochastic transportation-location problems, *Computers & Mathematics with Applications*, Vol 3, Pág. 87-94.
- [38]. Letitia M. Pohl, Russell D. Meller, Kevin R. Gue. (2009) An analysis of dual-command operations in common warehouse designs, *Transportation Research Part E: Logistics and Transportation Review*, Vol 45, Pág. 367-379.
- [39]. LF, Moreno Velasquez, JD, Velasquez Henao y FJ, Díaz Serna. (2005) Solución al problema combinado de ubicación estratégica de almacenes y asignación de inventarios usando técnicas heurísticas. Pág. 7257-66.
- [40]. Lu Chen, André Langevin, Diane Riopel, (2011) A tabu search algorithm for the relocation problem in a warehousing system, *International Journal of Production Economics*, Vol 129, Pág. 147-156.
- [41]. M.S. Hung, John C. Fisk, (1984) Economic sizing of warehouses: A linear programming approach, *Computers & Operations Research*, Vol 11, Pág. 13-18.
- [42]. Mier, Ortega, y J, Delgado. (2002) Utilización de metaheurísticas para el dimensionamiento y localización de instalaciones industriales. II Conferencia de Ingeniería de Organización. Pág. 623-621.
- [43]. O. Hounkonnou, (1990) Modelling and simulation of an automated warehouse in activity, *Mathematical and Computer Modelling*, Vol 14, Pág. 1100-1102.
- [44]. Peter Baker, Marco Canessa, (2009) Warehouse design: A structured approach, *European Journal of Operational Research*, Vol 193, Pág. 425-436.
- [45]. René de Koster, (2007) Tho Le-Duc, Kees Jan Roodbergen, Design and control of warehouse order picking: A literature review, *European Journal of Operational Research*, Vol 182, Pág. 481-501.
- [46]. Ria Kalfakakou, Stefanos Katsavounis, Kostas Tsouros, (2003) Minimum number of warehouses for storing simultaneously compatible products, *International Journal of Production Economics*, Pág. 559-564.
- [47]. Ronald H. Ballou, (1975). A multiproduct plant/warehouse location model with nonlinear inventory costs, *Journal of Operations Management*, Vol 5, Pág. 75-90.
- [48]. Roy, t.j. y Erlenkotter, d. (1982). A dual based procedure for dynamic facility location. *management science*. vol. 28 (10) 1091-1105.
- [49]. Sebastian Henn, (2012) Algorithms for on-line order batching in an order picking warehouse, *Computers & Operations Research*, Vol 39, Pág. 2549-2563.
- [50]. Sebastian Henn, Gerhard Wäscher, (2012). Tabu search heuristics for the order batching problem in manual order picking systems, *European Journal of Operational Research*, Vol 222, Pág. 484-494.
- [51]. Semih Önüt, Umut R. Tuzkaya, Bilgehan Doğanç, (2008). A particle swarm optimization algorithm for the multiple-level warehouse layout design problem, *Computers & Industrial Engineering*, Vol 54, Pág. 783-799
- [52]. Stefan Minner, (2009). A comparison of simple heuristics for multi-product dynamic demand lot-sizing with limited warehouse capacity, *International Journal of Production Economics*, Vol 118, Pág. 305-310
- [53]. Steffey, William. (2008). Layout design analysis for the storage area in the distribution center.



- [54]. T.C. Poon, K.L. Choy, Harry K.H. Chow, Henry C.W. Lau, Felix T.S. Chan, K.C. Ho, (2009). A RFID case-based logistics resource management system for managing order-picking operations in warehouses, *Expert Systems with Applications*, Vol 36, Pág. 8277-8301.
- [55]. Venkata Reddy Muppani (Muppant), Gajendra Kumar Adil, (2008). A branch and bound algorithm for class based storage location assignment, *European Journal of Operational Research*, Vol 189, Pág. 492-507
- [56]. Venkata Reddy Muppani (Muppant), Gajendra Kumar Adil, (2007). Efficient formation of storage classes for warehouse storage location assignment: A simulated annealing approach, *Omega*, Vol 36 Pág. 609-618
- [57]. W. Yang, Felix T.S. Chan, V. Kumar, (2012). Optimizing replenishment policies using Genetic Algorithm for single-warehouse multi-retailer system, *Expert Systems with Applications*, Vol 39, Pág. 3081-3086
- [58]. Y., Hinojosa, Y., Puerto, J., Fernández.(2000). Multiperiod two-echelon multicommodity capacitated plant location problem. *European Journal of Operational Research*, vol. 123, Pág. 271-291.
- [59]. Yon-Chun Chou, Yao-Hung Chen, Hui-Min Chen, (2012). Recency-based storage assignment and warehouse configuration for recurrent demands, *Computers & Industrial Engineering*, Vol 62, Pág. 880-889.

Ing. Andrés Mauricio Hualpa Zúñiga

Es Ingeniero Industrial de la Universidad Cooperativa de Colombia. Estudiante de Maestría en Ingeniería Industrial de la Universidad Distrital Francisco José de Caldas. Director del Programa de Ingeniería Industrial Universidad INCCA de Colombia. Docente Investigador Universidad Cooperativa de Colombia. Email: andresmhz@hotmail.com

Ing. Carolina Suárez Roldan

Es Ingeniera Industrial de la Universidad de Ibagué. Estudiante de Maestría en Ingeniería Industrial de la Universidad Distrital Francisco José de Caldas. Actualmente se desempeña como Coordinadora de Investigación Ingeniería Industrial Universidad Cooperativa de Colombia. Email: coutrs32@gmail.com

Gestión de la cadena de abastecimiento del biodiésel: una revisión de la literatura

Biodiesel supply chain management: A survey paper

Manuel Barón

Universidad Católica
de Colombia
Facultad de Ingeniería
mjbaron@ucatolica.edu.co

Isaac Huertas

Universidad Católica
de Colombia
Facultad de Ingeniería
ihuertas@ucatolica.edu.co

Javier Orjuela

Universidad Distrital
jaorjuela@etb.net.co



Resumen

Los beneficios asociados a los biocombustibles se ven reflejados en un creciente número de países, introduciendo o planeando introducir, políticas para incrementar la proporción de su producción dentro de su matriz energética. Actualmente, sólo pequeñas cantidades de biocombustibles se transan en los mercados internacionales ya que la mayoría se consume domésticamente. Sin embargo, se espera que el comercio de biocombustibles se expanda rápidamente dado que numerosos países no tendrán la capacidad doméstica para abastecer sus mercados internos.

Con base en el tendiente crecimiento del negocio de los biocombustibles, tanto a nivel mundial como local, resulta pertinente contar con herramientas técnicas que apoyen la toma de decisiones en la gestión agroindustrial para aprovechar de mejor manera un mercado por ahora naciente, pero que a mediano y largo plazo se escenifica como un gran sector productivo que podrá generar desarrollo a las regiones productivas. Este trabajo tiene el fin de presentar un estado del arte sobre la cadena de abastecimiento del biodiésel y los modelos matemáticos desarrollados para apoyar la toma de decisiones en la gestión de la cadena de suministro.

Palabras clave: biomasa, biocombustibles, biodiésel, gestión de la cadena de suministro.

Abstract

The benefits of biofuels have been reflected in an increasing number of countries that are introducing (or planning to introduce) policies to increase the proportion of biofuels in their energy mix. Currently, only small quantities of biofuels are traded on international markets because most of these fuels are consumed domestically. However, a rapid ex-

Fecha recibido: mar. 16/2013
Fecha modificado: may. 23/2013
Fecha aceptado: may. 31/2013



pansion of bio-fuel trade is expected since many countries have no domestic capacity to supply their own demand.

Based on the expected-growth trend of the bio-fuels business, both globally and locally, it seems appropriate to have the technical tools to support decision making in agribusiness management so as to make better use of a so-far nascent market; yet knowing that in the medium and long term, such a market may become a large manufacturing sector that brings development to its productive regions. This paper attempts to present a comprehensive survey of the bio-diesel supply chain together with mathematical models intended to support decision-making in supply chain management.

Key words: Biomass, Biofuels, Biodiesel, Supply Chain Management (SCM).

1. Introducción

Los combustibles de origen fósil han sido útiles en el desarrollo de la sociedad; en Colombia por ejemplo, han sido históricamente su mayor fuente de energía. Sin embargo, la diversificación de fuentes de energía primarias favorece la seguridad energética al disminuirse la dependencia de una sola fuente. Se hace importante fomentar la diversificación tecnológica para usos de combustibles diferentes a los tradicionales, principalmente la referente a fuentes renovables, tales como: hidroeléctrica, geotérmica, eólica, solar, mareomotriz, biomasa, biogás y cultivos energéticos.

El desarrollo de los sistemas de bioenergía requiere de la acción multidisciplinaria de diversos sectores involucrados en la cadena productiva: productos agrícolas y energéticos. Se requiere pues, plantear lineamientos y estrategias para el desarrollo de la industria del biocombustible biodiesel, lo cual depende de las condiciones locales específicas como son el clima, la disponibilidad de agua y las condiciones del suelo, infraestructura, capacidades de logística y procesamiento industrial de productos, subproductos y residuos orgánicos, necesidades del sector energía y aspectos agronómicos, económicos y sociales implicados en el desarrollo rural sustentable.

En los últimos años, los biocombustibles han comenzado a ser considerados mundialmente como una alternativa seria frente al petróleo [1]. Esto básicamente, obedece a dos razones: se estima que la disponibilidad de las reservas de petróleo llegará pronto a su fin (aproximadamente 100 años) y el precio del crudo, el cual tiende a aumentar.

Ante el panorama anterior, el mercado de los biocombustibles surge como alternativa viable, aunque se encuentra en sus primeros estadios y se localiza principalmente a nivel nacional o regional [2]. La producción, en países en vía de desarrollo (asiáticos y suramericanos) y la refinación concentrada en naciones tecnificadas industrialmente (Norteamérica y Europa del Este). En cuanto a los países de América Latina, estos presentan condiciones ideales para la producción de biocombustibles tales como el alto porcentaje de áreas húmedas (40%) y los recursos hídricos renovables, los bajos costos de producción

en las zonas tropicales o a sus bajos costos salariales [3]. La producción refinación del biodiésel es un área en progreso de la industria de los combustibles, que reclama logros tecnológicos debido a la relevancia y ventajas ambientales [4].

No obstante, un aspecto que todavía es debatido en el mundo, es si el balance energético del biodiésel es positivo entendido como la diferencia entre la energía que produce un kilogramo de combustible (biodiésel en este caso) y la energía necesaria para producirlo, lo cual incluye extracción (cultivo, en este caso), procesamiento, transporte y refinado de la materia prima, entre otros [5].

Este artículo se limita a la cadena productiva del biocombustible biodiesel obtenido a partir de la palma de aceite. Aquí se abordan elementos económicos, ambientales y energéticos pertinentes para avanzar en el desarrollo de una cadena productiva del biocombustible biodiésel. El análisis abarca la cadena productiva del biocombustible desde la producción de palma hasta su utilización como biodiesel, identificando temas que requieren atención por parte de los sectores privado y público.

El artículo se organiza de la siguiente forma. La Sección 2 explica la metodología utilizada para realizar la revisión. La Sección 3 describe el biodiésel como fuente de energía alternativa frente a los combustibles fósiles, sus diferentes formas de producción, sus ventajas y desventajas, su evolución y futuro y, finalmente, su impacto ambiental (uno de los temas más controversiales). La Sección 4 trata la Cadena Productiva del Biodiésel, considerando sus eslabones y estructura, y los indicadores de eficiencia respectivos. Finalmente, la Sección 5, hace una descripción de las herramientas utilizadas para la gestión de la cadena, apoyándose en modelos matemáticos de Programación Matemática, Dinámica de sistemas, Simulación y otros.

2. Metodología

Este documento inicia con la construcción del contexto del biodiesel, para lo cual se recurrió a la recolección, procesamiento y análisis de información secundaria proveniente de: bases de datos científicas (ScienceDirect, SpringerLink, EBSCO-HOST, Scielo y Scopus) y congresos (Congreso Latinoamericano de Dinámica de Sistemas); los Ministerios colombianos de Agricultura y del Medio Ambiente; entes nacionales como Fedebiocombustibles; entes internacionales como USDA (Departamento de Agricultura de Estados Unidos) y NACEPT (Consejo Consultivo Nacional de Políticas Ambientales y Tecnología) y organizaciones multinacionales como IICA (Instituto Interamericano de Cooperación para la Agricultura) y Cepal (Comisión Económica para América Latina y el Caribe). El consolidado de número de artículos por fuente se observa en la Figura 1. Finalmente se realizó la revisión, análisis y clasificación de modelos matemáticos de SCM (*Supply Chain Management*) para cadenas productivas de biodiesel y cadenas afines: bioetanol, y otras de biomasa (caña de azúcar, cultivos forestales, entre otros). La fuente de información utilizada fueron las bases de datos científicas enunciadas.

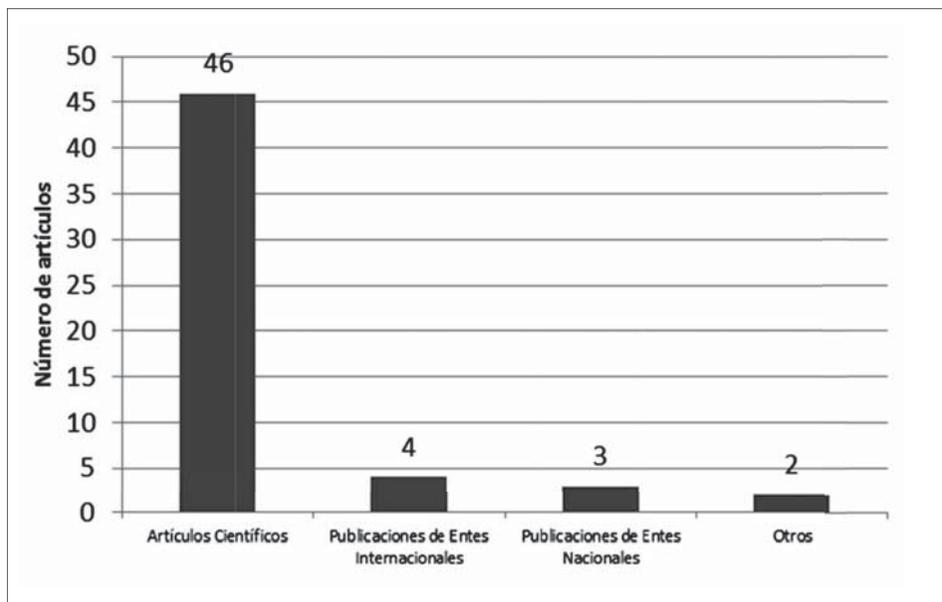


Figura 1. Distribución de información secundaria consultada.

Fuente: Los autores

3. Contexto del biodiésel

3.1. Definición

El biodiésel se ha definido como ésteres monoalquílicos de ácidos grasos de cadena larga, obtenidos a partir de lípidos renovables, tales como aceites vegetales o grasas animales, para su uso en motores por compresión (motores Diesel) [6] [7]. Se trata de un combustible biodegradable, cuyo uso disminuye las emisiones de gases de efecto invernadero y óxidos de azufre y la cantidad de hidrocarburos totales no quemados [5].

El biodiésel es actualmente puesto a la vanguardia como el combustible alternativo viable para los motores de combustión interna, puede ser producido a partir de fuentes renovables, a través de transesterificación, compatible con las infraestructuras tecnológicas existentes [8]. Puede ser usado en cualquier mezcla con el diesel obtenido de petróleo, debido a que tiene características similares, con menores emisiones de gases de escape [9].

Aunque potencialmente se pueden usar múltiples fuentes alternativas para obtenerlo, casi la totalidad del biodiésel que se produce actualmente proviene de los aceites de origen vegetal y, en particular, de tres cultivos: la palma, la soja y la colza (Figura 2). Estos tres cultivos proveen el 76% de la oferta mundial de aceite. El restante 24% se origina en cultivos como maní, algodón, oliva y girasol [5] [10].

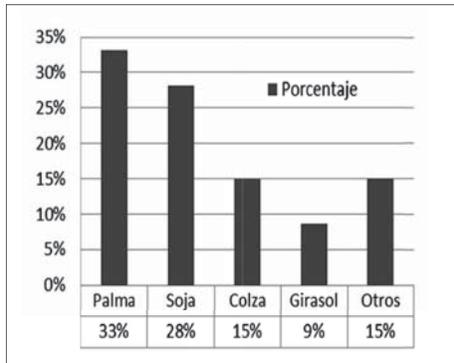


Figura 2. Composición de la oferta de aceite vegetal por fuente
Fuente: Los autores

3.2. Producción del biodiésel

Las principales etapas en la producción de los biocombustibles, dentro de las cuales se encuentra el biodiésel, se presentan en la Tabla I.

Según la Tabla I, el biodiesel puede provenir de plantas oleaginosas como Soja, Palma de aceite, Colza y Jatropha. Los procesos incluyen inicialmente el tratamiento de los cultivos, posteriormente un prensado o extracción del aceite y finalmente transesterificación. También,

MATERIAS PRIMAS			PROCESAMIENTO INDUSTRIAL
CULTIVOS ENERGÉTICOS		CULTIVO	PRE-TRATAMIENTO
PLANTAS OLEAGINOSAS	Soja	Cría (tratamientos genéticos) mejora del contenido en aceites, azúcar y almidón; búsqueda de variedades resistentes a la sequía; tecnología para mejora de los rendimientos	
	Palma aceitera		
	Colza		
	Jatropha		
CULTIVOS AZÚCAR/ALMIDÓN	Caña de azúcar		
	Sorgo dulce		
	Remolacha		
	Mandioca		
BIOMASA LIGNO-CELULÓSICA	Eucalipto	Cría (tratamientos genéticos) reducción del contenido de lignina; selección de las variedades con mayor adaptabilidad	Gasificación
	Sauce		Calentamiento y secado
	Switchgrass		Física: conminación mecánica, explosión a vapor, explosión de fibra de amonio, pirólisis; Química: hidrólisis ácida y alcalina; Biológica: reacción a microbios (hongos blancos, marrones y de pudrición blanda).
	Bermudagrass		
	Miscanthus		
RESIDUOS	Relleno sanitario		
	Desecho animal		

Fuente: A partir de Cepal [11]



el biodiesel se puede obtener de biomasa lignocelulósica pero con un proceso más complejo.

Las materias primas más utilizadas en los procesos productivos para la obtención de biocombustibles se relacionan en la Tabla II. Por su parte, las principales etapas usualmente incorporadas en la *BSC (Biofuels Supply Chain)* son los siguientes (ver Figura 3): producción de materias primas (que se relaciona con la disponibilidad de tierras y la idoneidad, la eficiencia del suelo asociado a diferentes tipos de plantas); la producción de biocombustibles (que se refiere a la transformación de las materias primas en biocombustibles a través de diversos procesos de conversión); mezcla (en el caso de que los biocombustibles se proporcionan a los consumidores finales se mezcla con los combustibles convencionales); el transporte de biocombustibles; y finalmente, el consumo en la red de distribución.

PROCESAMIENTO		PRODUCTO PRIMARIO		PROCESAMIENTO	PRODUCTO FINAL	
Prensado o extracción		Aceite vegetal		Transesterificación	Biodiesel	
					Bio-oil	
Molienda e hidrólisis		Azúcar		Fermentación y cofermentación	Purificación	Etanol, butanol
Pirólisis		reacciones gas - sólido	reacciones fase gas	Syngas	Purificación	Síntesis catalítica
						Gas natural sintético
						Diesel Fischer-Tropsch
						Dimetiléter
Digestión anaeróbica		Biogas			Water gas shift reaction	Hidrógeno
Pirólisis flash		Bio-oil		Tratamiento hídrico y refinado		Biodiesel
Licuefacción hidrotermal						
Hidrólisis		Azúcar		Fermentación y cofermentación	Purificación	Etanol, butanol
Digestión anaeróbica		Biogas				Biogás

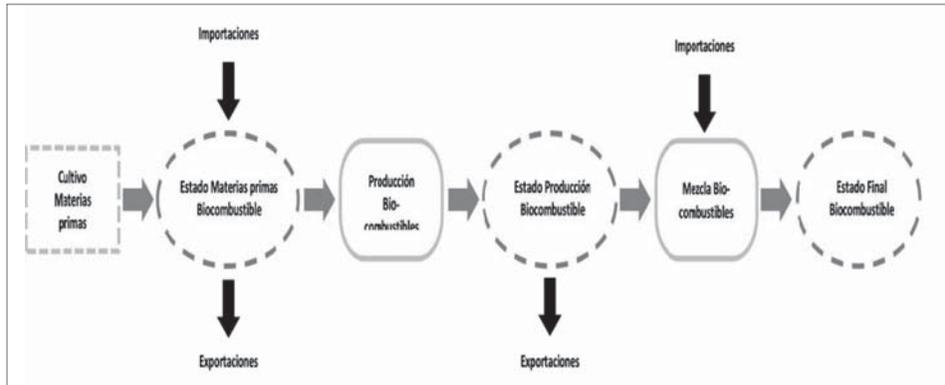


Figura 3. Etapas de la cadena de abastecimiento de los biocombustibles
Fuente: A partir de Papapostolou et al. [13]

Tabla II. Materia primas utilizadas en la producción de Biocombustibles	
Biodiesel	Bioetanol
Palma africana,	Caña de azúcar,
Cocotero, Higuerilla,	Remolacha azucarera,
Aguacate,	Sorgo dulce, Maíz, Yuca
Jatropha/Piñón,	
Colza/Canola, Maní,	
Soya, Girasol	

Fuente: A partir de Orjuela et al. [12]

El biodiésel puede ser producido económicamente en un amplio rango de lugares, tanto de asentamiento rural como urbano, y en diferentes escalas dependiendo de los requerimientos específicos: pequeñas para autoconsumo o comerciales para usos industriales [5]. En consecuencia, el biodiésel es actualmente considerado como el combustible alternativo más viable para los motores [8].

Los costos de producción para el biodiésel, varían ampliamente de un proceso a otro y de una región a otra: estas diferencias se deben, fundamentalmente, a costos de las materias primas, tipo de energía usado para el procesamiento, tanto calórica como eléctrica, a precios obtenidos para los subproductos derivados durante el proceso de producción [2]. La estructura del costo de producción del biodiesel se presenta en la Figura 4.

Ng et al. [8] sugiere que las estrategias requeridas para una industria de biodiésel estable y sostenida deberán estar basadas principalmente en los principios del libre mercado con un mínimo de “intervenciones artificiales”, por lo que hacen necesarias las políticas pertinentes y los avances tecnológicos apropiados en técnicas de producción y obtención de materias primas para mantener su competitividad económica.

3.3. Ventajas

Según Mittelbach y Remschmidt [14] y Knothe et al. [15], las principales ventajas del biodiésel son su potencial para reducir la dependencia de una determinada economía en el

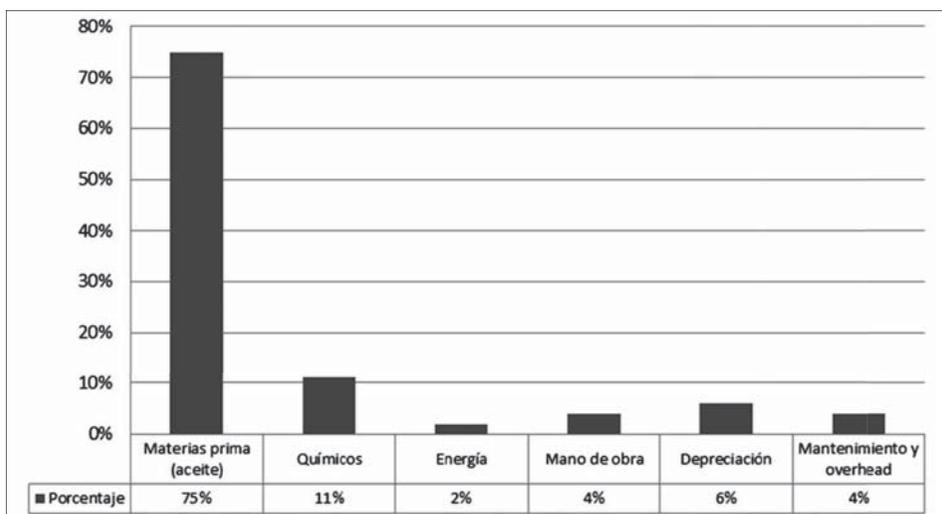


Figura 4. Estructura típica de los costos de producción del biodiésel.
Fuente: Esta investigación, Elaborada a partir de datos reportados en IICA [5]

petróleo, biodegradabilidad, alto punto de inflamación y lubricidad. Por su parte, Knothe et al. [15] proponen otras ventajas del biodiésel tales como: portabilidad, facilidad de obtención, renovación, mayor eficiencia de la combustión, menores contenidos de azufre y compuestos aromáticos.

Para Demirbas [16], la mayor ventaja que el biodiésel tiene sobre el diésel del petróleo es su respeto al medio ambiente. Trae beneficios económicos que incluirían el valor agregado a la materia prima, un mayor número de empleos en la manufactura rural, un aumento de impuestos sobre la renta y la inversión en maquinaria y equipo. El biodiésel puede ofrecer otros beneficios, incluyendo la reducción de las emisiones de gases de efecto invernadero, el desarrollo regional y la estructura social, especialmente a los países en desarrollo [17].

Demirbas [9] resalta que el biodiésel tiene mejores propiedades que el combustible diésel de petróleo; es renovable, biodegradable, no tóxico y esencialmente libre de azufre y compuestos aromáticos.

En igual sentido, Yee et al. [18] sostiene que la producción de biodiésel de palma y de colza no trae consecuencias negativas para el medio ambiente, ya que la cantidad de CO₂ emitido a la atmósfera es mucho menor que el CO₂ absorbido por esta. Adicionalmente, argumenta que la combustión de biodiésel de palma es más favorable al medio ambiente que la de los derivados del petróleo, con una reducción significativa del 38% de las emisiones de CO₂, por cada litro consumido.

Por su parte Fazal et al. [19] en diferentes investigaciones de laboratorio, encontró que: el biodiésel proporciona mejor lubricación que el de combustible fósil diésel; puede

mejorar la combustión y, por lo tanto, tiene un mayor rendimiento térmico que el diésel proveniente del petróleo. La potencia del motor se reduce un poco o nada en absoluto, porque el consumo de biodiésel aumenta lo suficiente para compensar su menor valor de calefacción. No obstante, en general, el biodiésel permite un rendimiento aceptable del motor y podría mejorarse aún más si la viscosidad puede ser reducida.

Por último, Chouinard et al. [20] afirma que el total de gases de efecto invernadero (GEI) de los biocombustibles son reducidos en comparación con los de combustibles fósiles, por absorción de dióxido de carbono por la biomasa durante el crecimiento de las plantas. Por otra parte, mientras que las reservas de combustibles fósiles están disminuyendo y la demanda mundial de energía va en aumento, la biomasa pasa a ser un recurso renovable.

3.4. Desventajas

Según Bala [21], el biodiésel es más caro de producir que el diésel de petróleo, lo que parece ser el factor principal en la no generalización de su uso. La producción actual mundial de aceite vegetal y grasa animal no es suficiente para reemplazar el uso de combustibles líquidos fósiles.

Igualmente, para Ng et al. [8], el biodiésel todavía no es económicamente viable para la adopción a gran escala en el día de hoy, debido principalmente al alto costo de las materias primas convencionales. Estudios de factibilidad realizados han demostrado que el biodiésel no es tan competitivo económicamente como el diésel fósil sin la intervención de los responsables políticos.

Respecto a la producción del biodiesel, Demirbas [9] sostiene que en el proceso se libera glicerina, la cual todavía constituye un problema por su contenido tóxico (aunque éste es moderado) y contaminante. Según IICA [5] el biodiesel aporta un 10% más de óxidos nitrosos a la atmósfera (causantes de la lluvia ácida) con respecto a los combustibles fósiles.

3.5. Evolución y futuro del biodiesel

3.5.1. En el contexto mundial

La producción mundial de biodiésel se mantuvo relativamente estable entre dos y tres millones de toneladas anuales hasta el 2004 (Figura 5). En el 2005 la producción aumenta alcanzando un nivel de 20 mil millones de litros en el 2010 [11] [5].

A pesar de que en la actualidad los biocombustibles representan poco más del 1% de la demanda total de combustible para transporte (y entre el 4% y el 7% para

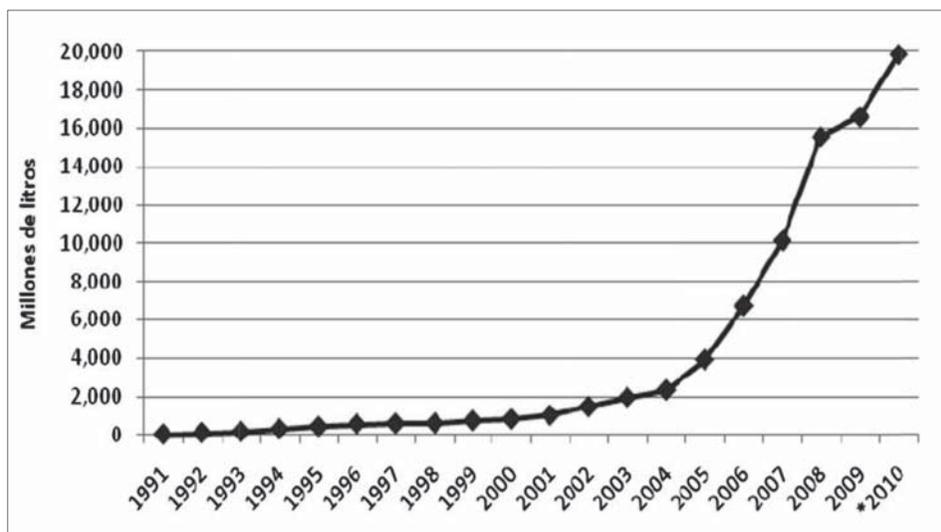


Figura 5. Producción de biodiésel en el mundo desde el año 1991 hasta el 2010.

Fuente: Gráfico obtenido de Cepal [11]

el 2030 según las proyecciones de la Agencia Internacional de Energía, AIE), los bio-combustibles parecen una buena oportunidad frente a los elevados precios del crudo internacional [2].

Hoy en día los costos de producción del biodiésel son entre 1, 5 y 3 veces más altos que para el diésel fósil. No obstante, cuando esto se haya atenuado, el biodiésel será un combustible de motor razonablemente disponible en un futuro cercano [9]. Igualmente, debido al incremento de los precios de la gasolina y el diésel fósil en todos los escenarios internacionales, y teniendo en cuenta el resto de exenciones fiscales, los biocombustibles se volverán competitivos en próximos años [22].

El reciente aumento en el uso potencial de biodiésel se debe no sólo al incremento del número de plantas productoras, sino también al tamaño de las instalaciones utilizadas en su producción. El crecimiento en la industria del biodiésel se espera que tenga un impacto significativo en el precio de las materias primas del biodiésel. Este crecimiento en la industria del biodiésel aumentará la competencia [9].

Con la inminente disponibilidad comercial de la futura generación de biodiésel, económicamente cada vez más factible, este combustible debe permanecer como un sustituto del diésel fósil que potencialmente puede satisfacer a mediano plazo las necesidades de energía para formar parte de la solución total a la energía en todo el mundo y los problemas ambientales [8]. Según Bala [21], el biodiésel es más caro de producir que el diésel de petróleo, lo que parece ser el factor principal en la no generalización de su uso. La producción actual mundial de aceite vegetal y grasa animal no es suficiente para reemplazar el uso de combustibles líquidos fósiles.

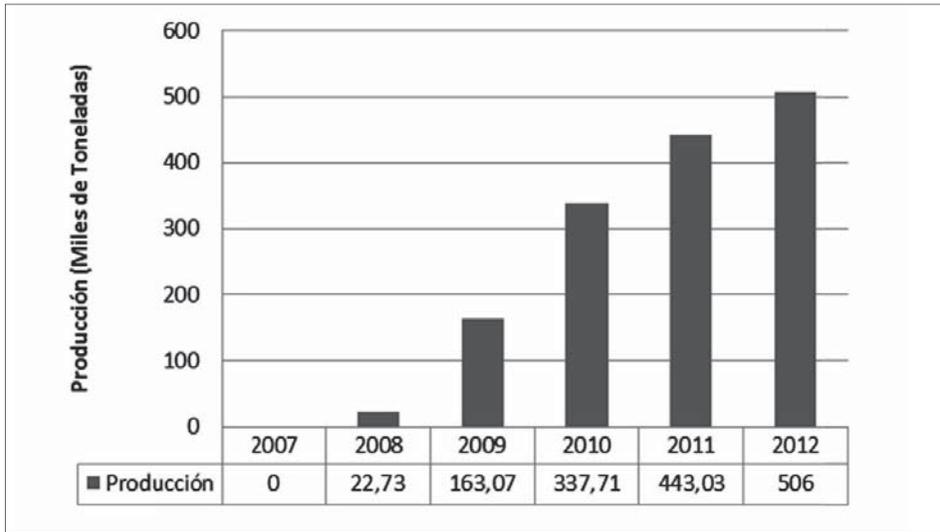


Figura 6. Producción de Biodiesel en Colombia, desde el año 2007 hasta el 2012.

Fuente: Esta investigación. Elaborado a partir de datos reportados en Fedebiocombustibles [23]

3.5.2. En Colombia

La producción de biodiesel en Colombia se presenta en la Figura 6 [23]. Como se observa, la producción de Biodiesel comienza en el año 2008 y su nivel ha aumentado hasta superar las 440.000 toneladas en el año 2011. La producción para el año 2012 se proyecta considerando la producción hasta el mes de Julio donde se alcanza un total de 282.600 toneladas.

En 2011 el biodiesel se constituyó en el principal mercado local para el aceite de palma. Las plantas de refinación de biodiesel y sus capacidades se presentan en la Tabla III [23] [24].

Adicionalmente existen dos plantas de una producción mucho menor: una en Tuma-co (Nariño), con una capacidad de 3.000 litros/día, y otra en Zulia (Norte de Santander) con una capacidad de 20.000 litros/día.

En Colombia se observa una tendencia favorable para los productores de biocombustibles en cuanto a legislación y políticas [25]. Colombia regula mediante la ley 939 de 2004 la producción de biodiesel y bioetanol los cuales deben ser usados en una mezcla con gasolina que irá aumentando a través de los años [3].

También es relevante la clara intención gubernamental por ganar mercado y competitividad en el campo internacional. Sin embargo, se evidencian los conflictos entre las políticas agrícolas y las de biocombustibles [25].

**Tabla III.** Capacidad de producción de Biodiesel en Colombia.

Empresa	Región	Capacidad (toneladas/año)	Capacidad (litros/día)	Fecha entrada en operación
Oleoflores	Codazzi, Cesar	70.000	169.000	Enero 2008
Odin Energy	Santa Marta, Magdalena	36.000	121.000	Junio 2008
Biocombustibles sostenibles del Caribe	Santa Marta, Magdalena	100.000	337.000	Marzo 2009
Bio D	Facatativá, C/marca	100.000	337.000	Febrero 2009
Manuelita	San Carlos de Guaroa, Meta	100.000	337.000	Julio 2009
Clean Energy	Barranquilla, Atlántico	40.000	134.976	Octubre 2009
Ecodiesel	Barrancabermeja, Santander	100.000	337.000	Junio 2010
Biocastilla	Castila la Nueva, Meta	10.000	33.744	Diciembre 2009
Total		556.000	1.806.720	

Fuente: Esta investigación. Elaborada a partir de datos reportados en Fedebiocombustible [23] y MinAgricultura [24]

3.6. Aspecto Ambiental

Los biocombustibles podrían brindar mayores ganancias económicas, desarrollo rural, reducción de la pobreza y menores emisiones de gases con efecto invernadero respecto de los combustibles fósiles. La producción de cultivos energéticos podría llevar a una expansión de la frontera agrícola, deforestación, monocultivos, contaminación del agua, difusión de organismos genéticamente modificados, problemas sobre la seguridad alimentaria y un empeoramiento de las condiciones laborales, son entre otras las preocupaciones de su crecimiento [1].

Contrario a algunos reportes que cuestionan la sostenibilidad del aceite obtenido de palma africana como fuente ecológica de energía, los resultados de este estudio del Análisis del Ciclo de Vida, ACV, han demostrado que el diésel de palma tiene el potencial para convertirse en la principal energía renovable en el futuro, con una relación de energía positiva y una reducción significativa de las emisiones de CO₂. Una investigación encontró que la utilización de biodiésel de palma podría generar una tasa de rendimiento de energía de 3.53 (la producción de energía / energía de entrada), lo que indica una energía positiva neta [18].

Sin embargo, Dufey [1] sostiene que los beneficios y costos de los biocombustibles varían ampliamente, de acuerdo al tipo de cultivo energético, método de cultivo, tecnología de conversión y las condiciones naturales de cada país o subregión. El uso de biodiésel debe permitir un equilibrio que debe buscarse entre la agricultura, el desarrollo económico y el medio ambiente [16].

Conviene resaltar que las políticas relativas a la producción de biocombustibles requieren ser analizadas en Colombia, así como en otros países de América del Sur, debido

a que la población de este continente presenta altos niveles de hambre y la utilización de suelo para los dos propósitos pondría en riesgo la seguridad alimentaria [3].

4. Cadena productiva del biodiesel

Gualteros [26] a partir de la información recopilada de la cadena productiva del biodiésel en diferentes países y del estado actual de la cadena en Colombia, propone un esquema de la cadena productiva del biodiésel a partir de palma africana, en el cual se incluyen no solo los eslabones de producción de biodiésel sino el entorno institucional y organizacional de la cadena, con el fin de tener una visión global del proceso y poder, además, identificar los posibles actores de la cadena que puedan participar en el estudio de la misma.

La Figura 7 describe la estructura detallada de la cadena. En ella se destacan los diferentes eslabones pertenecientes al proceso productivo, a la comercialización del biodiésel y a sus interacciones. Como se puede apreciar, la cadena productiva se divide en tres grandes sectores: el Agrícola, el Industrial y el de Servicios; y cada uno de ellos se subdivide en otros eslabones los cuales conforman la totalidad de la cadena.

Según Gualteros [26], en Colombia las materias primas de biomasa son transportadas generalmente en camiones desde las fincas a la planta de refinación de biocombustibles. Las instalaciones de almacenamiento son necesarias entre las fincas y las biorrefinerías. El almacenamiento de pre-tratamiento también se proporciona para garantizar la frescura de las materias primas y aumentar la tasa de rendimiento. En la mayoría de los casos, la materia prima o las materias primas son transportadas desde las fincas directamente a la refinería. Las materias primas de biomasa se convierten en productos terminados. Una breve descripción del funcionamiento de los elementos o subsistemas de la cadena de suministro de biocombustibles se presenta en la Tabla IV [27].

Según IICA [5], se reconocen cuatro grandes sectores interrelacionados en la cadena productiva de biodiésel: 1) Un sector productor del aceite, materia prima fundamen-



Figura 7. Esquema general de la cadena productiva.
Fuente: A partir de [26]



Tabla IV. Principales sistemas en la cadena de abastecimiento de los Biocombustibles	
Sistema de la cadena de abastecimiento	Descripción de las actividades
Producción de mate-rias primas	De biomasa herbácea y de madera disponible regionalmente, incluidos los residuos agrícolas y forestales, cultivos energéti-cos y materiales de desecho.
Logística de materias primas	Implementar la infraestructura para biomasa, equipos, mano de obra, y los sistemas para llevar a cabo la cosecha, recolección, almacena-miento, pre-procesamiento, y las operaciones de transporte.
Producción de Bio-combustibles	Implementar de forma rentable las instalaciones de conversión de la biomasa a biocombustibles, cumpliendo con todos los reglamentos de seguridad y medio ambiente para producir biocombustibles que cumplan con los estándares de calidad.
Distribución de los biocombustibles	Implementar la infraestructura de distribución de biocombus-tibles para llevar a cabo el almacenamiento, mezcla, transporte, y las opera-ciones de dispensación.
Uso final de los biocombustibles	Facilitar la disposición al público de los biocombustibles con un rendi-miento similar al diesel derivado del petróleo.

Fuente: A partir de NACEPT [27]

tal para la obtención del biodiésel, el cual comprende dos subcomponentes: un subsector de producción de las fuentes, vegetales y animales, de aceite, y otro subsector que transforma la fuente en aceite. 2) Un sector de producción de biodiésel, en el cual se procesa el aceite para la producción de biodiésel. 3) Un sector de mercado de biodiésel, donde se documenta el consumo aparente de biodiésel a nivel nacional, definido como consumo + (importaciones -exportaciones). Y 4) El marco legal y regulatorio, fundamental para el funcionamiento de la industria de producción y consumo sostenible de biodiésel.

Awudu y Zhang [28] proponen, como se observa en la Figura 8, el marco general de la cadena de suministro de biocombustibles. Los principales elementos de la cadena de suministro de biocombustibles que se consideraron son los siguientes: (1) granjas o fincas, (2) instalaciones de almacenamiento, (3) plantas de biorrefinería, (4) instalaciones de mezcla, (5) puntos de venta y el transporte (estación).

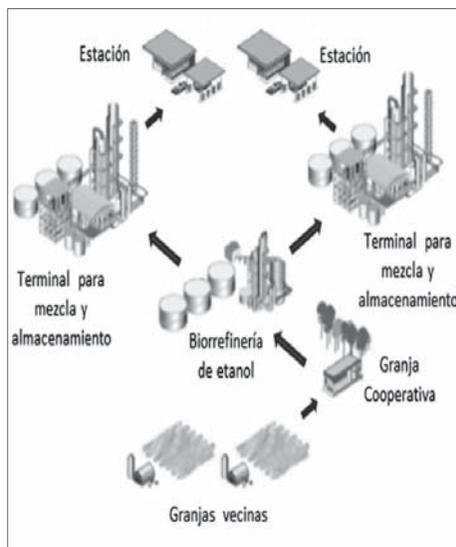


Figura 8. Cadena de suministro de biocombustibles según Awudu y Zhang.

Fuente: A partir de [28]

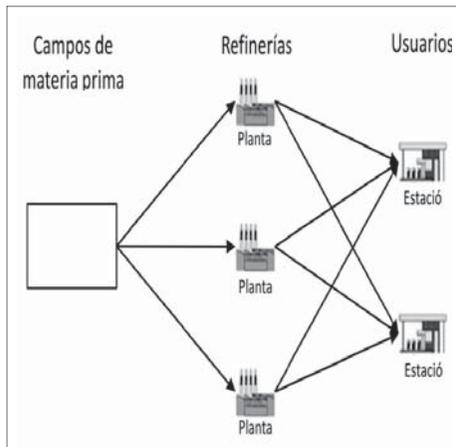


Figura 9. Cadena productiva de biocombustibles según Huang [29].
Fuente: A partir de Huang [29]

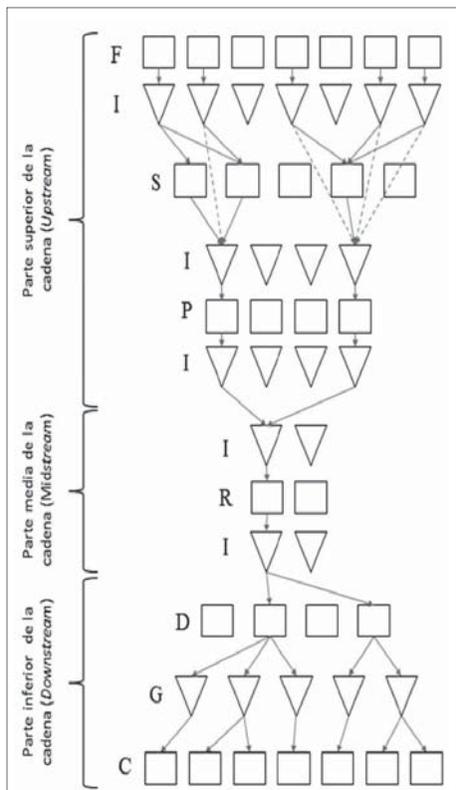


Figura 10. Estructura general de la cadena de biocombustible según An et al.
Fuente: A partir de [30]

De una manera más simplificada, Huang et al. [29] representa la cadena productiva de biocombustibles como una típica cadena de suministro. Por lo tanto, el problema de diseño de sistemas de biocombustibles lo considera dentro de la categoría general de los problemas de localización de la cadena de suministro (Figura 9).

Para An et al. [30], la estructura de flujo de red basada en materias primas de biomasa agrícola se muestra en la Figura 10. Esta red multinivel muestra las potenciales ubicaciones y capacidades de las fincas (F) y sitios de almacenamiento de la biomasa (S); pre-instalaciones de procesamiento (P), biorrefinerías (R), centros de distribución (D); y estaciones de servicio (G), que abastecen a los clientes (C). Los posibles lugares para los inventarios se denotan por I (Ss, Ds y Gs son los inventarios por naturaleza).

5. Herramientas matemáticas utilizadas para la Gestión de la cadena productiva del biodiesel

5.1 Clasificación

An et al. [30] en su artículo de revisión sobre modelos aplicados a la cadena productiva de biocombustibles y combustibles fósiles, concluyen que los estudios sobre biodiésel se relacionan principalmente con los procesos en los niveles superiores de la cadena, o los procesos “*upstream*”, (es decir, de los proveedores de biomasa a las plantas de conversión) especialmente a niveles de decisión operacional o integrados

**Tabla V.** Resumen artículos de SCM de biocombustibles

N°	Autor	Año	País	Herramienta
1	Chien y Yueyue	2011	USA	MISP
2	Kim et.al.	2011	USA y Corea	MILP Dinámica de sistemas
3	Orjuela et. al.	2011a	Colombia	Análisis multivariado Redes neuronales
4	Orjuela et. al.	2011b	Colombia	Dinámica de Sistemas
5	Papapostolou et. al.	2011	Grecia	MILP
6	Zhang et.al.	2011	China y Portugal	MILP
7	Zhu et. al.	2011	USA	MILP
8	Eksioglu y Petrolia	2010	USA	MILP
9	Flórez	2010	Colombia	Dinámica de sistemas
10	Huang et. al.	2010	USA	MILP
11	Zapata et. al.	2010	Colombia	Dinámica de sistemas
12	Eksioglu et. al.	2009	USA	MILP
13	Constantino, et. al	2008	Portugal	Programación entera
14	Lejars et. al.	2008	Francia	Paquete de simulación MAGI
15	Ravula et.al.	2008	USA	Simulación discreta
16	Dunnett et. al.	2007	Inglaterra	MILP
17	Kumar y Sokhansanj	2007	USA y Canadá	Modelo IBSAL
18	Sokhansanj et. al.	2006	USA y Canadá	Modelo IBSAL
19	Martins et. al.	2005	Portugal	MILP, column generation
20	Troncoso y Garrido	2005	Chile	MILP
21	Gunnarsson et. al.	2004	Suecia	MILP
22	Higgins y Postma	2004	Australia	MILP
23	Gigler et. al.	2002	Holanda	Programación dinámica (DP)
24	De Mol et. al.	1997	Holanda	Simulación (Prosim) Programación entera mixta (MIP)
25	Cundiff	1996	USA	Programación lineal
26	Gallis	1996	Grecia	Simulación

MISP: programación estocástica entera mixta.

MILP: Programación lineal entera mixta.

IBSAL: Logística y análisis de la Cadena de Biomasa Integrada (Integrated Biomass Supply Analysis and Logistics)

Fuente: Los autores

(de estratégico a operacional). La razón es que inicialmente las investigaciones se han enfocado en evaluar varios tipos de materias primas y probar la eficiencia de la logística del mismo.

En la revisión de la literatura no se encontraron estudios que optimicen las operaciones de conversión, en los niveles medios de la cadena o “*midstream*”, porque muchas

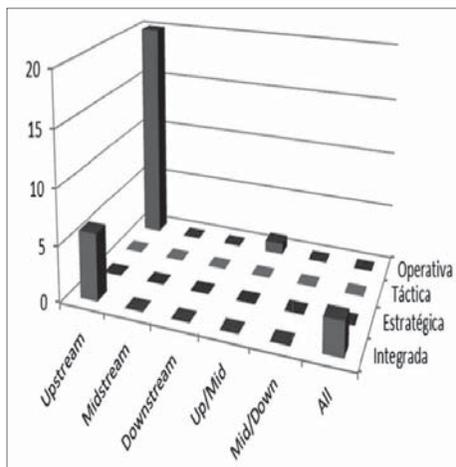


Figura 11. Categorización de los artículos relacionados a SCM de biocombustibles.

Fuente: A partir de [30]

características semejantes a las de la cadena del biodiesel (bioetanol y biomasa en general), hasta abril de 2012. Los artículos, país y herramienta matemática utilizada se recopilan en la Tabla V.

En la medida de los años recientes se ha notado un incremento en el número de artículos escritos con modelos matemáticos para SCM de biocombustibles y afines (Biomasa, Bioetanol, entre otras), tal como se presenta en la Tabla VI. En ella se observa que el país con mayor cantidad de estudios es USA, con un total de 10 artículos (38% del total de artículos revisados), mientras que la mayoría de los artículos restantes son de países de Europa (38%). El porcentaje restante (24%) se distribuye entre el Pacífico Sur y Suramérica. Algunos estudios fueron realizados colaborativamente entre dos países: USA - Canadá, USA - Corea del sur y Portugal - China. (En este caso se han asignado al país con más artículos). En los dos últimos casos, se observa una participación de países asiáticos como China y Corea del sur.

Dentro de los artículos revisados, nueve van destinados a cadenas de Biomasa Lignocelulósica y ocho a biomasa de residuos. Estas biomásas pueden ser utilizadas para la obtención de biodiesel, bioetanol, generación de calor u otros usos. Cinco artículos estudian la biomasa procedente de cultivos de azúcar y almidón, para la producción de bioetanol y alimento. Finalmente, cuatro artículos se refieren a biomasa de plantas oleaginosas para la producción de biodiesel y alimento. En la Figura 12 se representa el porcentaje correspondiente a las cadenas tratadas en los artículos.

En la Tabla VII se observa que la herramienta más utilizada para la formulación de modelos matemáticos es la programación matemática, y dentro de esta, la más utilizada es

tecnologías están todavía siendo desarrolladas y evaluadas. Adicionalmente, los autores afirman que varios estudios se han basado en un modelo de programación lineal entera mixta (MILP por sus siglas en inglés) de producción / distribución.

Una clasificación de los 30 artículos revisados por An et al. [30] se muestra en la Figura 11. En la figura se clasifican los modelos matemáticos según el nivel de la cadena productiva (*upstream*, *midstream* y *downstream*) y nivel de decisión (estratégico, táctico, operativo e integral).

En el presente estudio se revisaron 26 artículos sobre modelos matemáticos para apoyar la toma de decisiones en cadenas de biodiesel, y cadenas productivas con características semejantes a las de la cadena del biodiesel (bioetanol y biomasa en general), hasta abril de 2012. Los artículos, país y herramienta matemática utilizada se recopilan en la Tabla V.



Tabla VI. Clasificación de los artículos por país y periodo					
Países	1996-1999	2000-2003	2004-2007	2008-2011	Total
Australia			1		1
Colombia				4	4
Chile			1		1
Francia				1	1
Grecia	1			1	2
Holanda	1	1			2
Inglaterra			1		1
Portugal			1	2	3
Suecia			1		1
USA	1		2	7	10

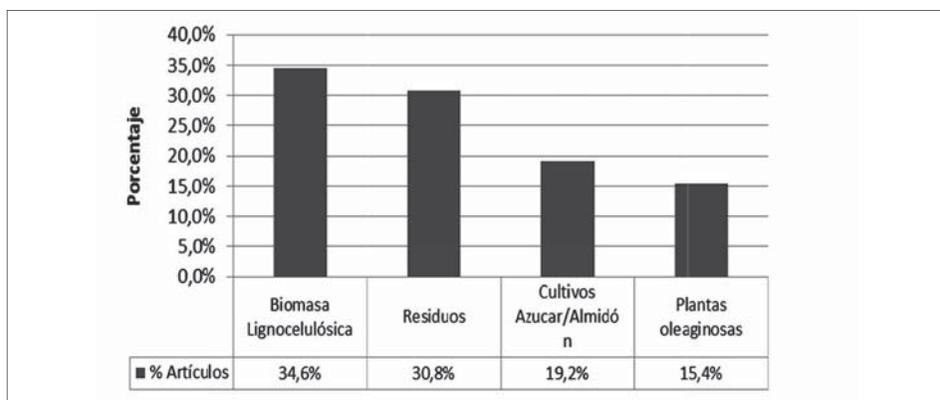


Figura 12. Distribución de los artículos de Supply Chain Management por cadena.
Fuente: Los autores

Tabla VII. Objetivos y Herramientas utilizados en los modelos de SCM								
Objetivo	Herramienta							Total
	DP	IBSAL	LP	MILP	MISP	Simulación	DS	
Diseño		2		1				3
Diseño y gestión				7	1	1		9
Gestión			1	2				3
Programación cosechas				3		1		4
Otros	1					2	4	6
Total	1	2	1	13	1	4	4	

DP: Programación Dinámica

IBSAL: Logística y análisis de la Cadena de Biomasa Integrada (Integrated Biomass Supply Analysis and Logistics)

LP: Programación lineal

MILP: Programación lineal entera mixta

MISP: Programación estocástica entera mixta

DS: Dinámica de Sistemas

Fuente: Los autores

la programación lineal entera mixta (MILP) (con un 50% de los artículos revisados). En segunda instancia, la herramienta más utilizada es la simulación y dinámica de sistemas (31%), teniendo en cuenta que se han desarrollado en diferentes aplicaciones computacionales.

El diseño y gestión de la cadena es el objetivo utilizado con mayor frecuencia en los modelos de SCM evaluados (35% del total de artículos revisados). En segunda instancia se encuentran modelos con objetivos varios (Otros) los cuales corresponden a análisis y cuantificación de variables en las cadenas productivas, evaluadas principalmente por las herramientas simulación y dinámica de sistemas.

Los objetivos de los modelos matemáticos encontrados se presentan detalladamente en la Tabla VIII. Se observa que la mayoría de los modelos se dirigen a optimizar los costos o ganancia de la cadena, independientemente si el modelo busca el diseño, gestión o programación de la cadena.

Tabla VIII. Objetivos y características de los modelos desarrollados en los artículos de SCM		
Autor	Año	Objetivo
Chien y Yueyue	2012	Diseño y gestión de la cadena del bioetanol Minimización los costos de la cadena
Kim et.al.	2011	Diseño y gestión de la cadena de biocombustibles Maximización ganancia total de la cadena
Orjuela et. al.	2011a	Pronostico del comportamiento de la producción de bioetanol Identificación de factores y variables relevantes en el sistema
Orjuela et. al.	2011b	Identificación y análisis de factores y variables relevantes en lo que respecta a la problemática de la dicotomía en el uso del suelo.
Papapostolou et. al.	2011	Diseño y gestión de la cadena de biocombustibles Maximización el rendimiento total de la cadena
Zhang et.al.	2011	Gestión de la cadena forestal Maximización el beneficio económico de los boques Minimización el impacto negativo en el hábitat forestal
Zhu et. al.	2011	Gestión de la cadena de biomasa Maximización ganancia
Eksioglu y Petrolia	2010	Diseño y gestión de la cadena de biocombustibles Minimización del costo de la cadena
Flórez et. al.	2010	Análisis de políticas actuales de producción de biocombustibles y de cambios posibles para alcanzar la producción deseada.



Huang et. al.	2010	Gestión de la cadena de suministro en un horizonte de planeación Minimización el costo de la cadena
Zapata et. al.	2010	Búsqueda de políticas que contribuyan al aumento de la oferta de biodiesel proveniente de palma de aceite.
Eksioglu et. al.	2009	Diseño de la cadena de suministro y gestión de una biorrefinería Minimización el costo de la cadena
Constantino, et. al	2008	Secuenciación de la cosecha
Lejars et. al.	2008	Programación de cosechas de la caña de azúcar Maximización ganancia
Ravula et.al.	2008	Disminución el tiempo de transporte del sistema
Dunnett et. al.	2007	Diseño de la cadena de Biomasa Minimización el costo total de la cadena
Kumar y Sokhansanj	2007	Diseño de la cadena de Biomasa Asegurar una producción neta de energía con una huella de carbono aceptable
Sokhansanj et. al.	2006	Diseño de un modelo operacional para el abastecimiento y transporte de la biomasa que cuantifique los recursos necesarios y el costo de operación
Martins et. al.	2005	Programación de cosechas Maximización el volumen de madera cortada
Troncoso y Garrido	2005	Diseño y gestión de la cadena Minimización el valor presente del costo de la cadena
Gunnarsson et. al.	2004	Gestión de la cadena Minimización costo total
Higgins y Postma	2004	Secuenciación de las cosechas Maximización rentabilidad
Gigler et. al.	2002	Ruteo y asignación de actividades Minimización el costo total de la cadena
De Mol et. al.	1997	Diseño de la cadena Minimización coto de la cadena Evaluación de costos y consumo de energía
Cundiff	1996	Gestión de la cadena de suministro Minimización los costos de la cadena
Gallis	1996	Evaluación de los efectos de variables logísticas sobre el costo

Fuente: Los autores

Las características y variables de decisión de cada uno de los modelos evaluados se relacionan en la Tabla IX. Se encuentra que la mayoría de modelos incluyen variables relacionadas con el flujo de mercancía (materias primas, productos intermedios y productos terminados) a través de la red, y variables relacionadas con la ubicación y tamaño (capacidad) de las instalaciones.

Tabla IX. Características y variables de los modelos matemáticos			
Autor	Año	Características	Variables
Chien y Yueue	2012	El modelo busca definir: tamaño y lugares de las biorrefinerías y las terminales, estrategia de asignación de recursos como materia prima, plan de producción de etanol, y plan de transporte de la materia prima y del combustible.	Localización de las biorrefinerías Localización de las terminales Capacidad (galones) de las biorrefinerías Galones de etanol producidos en las biorrefinerías Toneladas de Materias primas enviadas de las fuentes a las biorrefinerías Galones transportados de las biorrefinerías a las terminales Galones transportados de las terminales a las ciudades Déficit de galones en las ciudades
Kim et.al.	2011	El modelo se aplica para diseñar un sistema de refinamiento distribuido y uno más centralizado. El estudio evalúa los beneficios totales, valores, costo, y los diseños de la red de suministro de ambos sistemas.	Número óptimo, localización y tamaño de varios tipos de plantas de procesamiento La cantidad de biomasa, productos intermedios y producto final a ser transportado entre las locaciones en cada periodo de tiempo
Orjuela et. al.	2011 ^a	El modelo busca la predicción del comportamiento futuro de la producción de caña panelera. La dinámica de la cadena productiva frente a la problemática es estudiada a partir de la identificación de factores y variables relevantes y su diagrama causal.	Producción Rendimiento Participación de la caña panelera en la valoración de la producción Participación del número de personas en el sector Variación y superficie cosechada
Orjuela et. al.	2011 ^b	Con el modelo se analiza la problemática de la producción de bioetanol a partir de la caña panelera y la repercusión que tiene en la seguridad alimentaria de la población colombiana.	Sector Población Empleo Parque Automotor Combustible Panela y Siembra
Papapostolou et. al.	2011	Facilita y apoya la toma de decisiones en la planificación de diversas cuestiones operacionales, tales como las inversiones en infraestructura, las cantidades de materias primas que se cultivan, las cantidades de biocombustibles que se producen en el mercado nacional o importados, identificando la mejor	Cantidad de producción de cada cultivador en cada escenario de clima Cantidad transportada de cada cultivador a la planta central, en cada momento de tiempo, proveniente de cada tipo de bodega, en cada escenario de clima Capacidad de cada tipo de almacenaje, en cada periodo de tiempo, de cada cultivador, en cada escenario de clima



Papapostolou et. al.	2011	importados, identificando la mejor solución disponible para el diseño óptimo y el funcionamiento de la cadena de suministro de biocombustibles	Exceso de capacidad de almacenaje a la intemperie, en cada periodo de tiempo, de cada cultivador, en cada escenario de clima. Déficit en la biomasa enviada a la planta central en cada periodo, en cada escenario de clima.
Zhang et. al.	2011	El modelo considera la maximización del valor presente neto de los bosques, sujeto a restricciones de área de explotación total en cada periodo, y en la sostenibilidad de los flujos de madera aserrada.	Selección (si o no) de los sitios de recolección, planta de energía, sitios de pretratamiento Cantidad de flujo en la red
Zhu et. al.	2011	Sistema logístico de biomasa para la producción de biocombustible. Abarca desde la siembra y la cosecha hasta la entrega a la biorrefinería, incluyendo el manejo de residuos, centrándose en la integración de las decisiones estratégicas en el diseño de la cadena de suministro y las decisiones tácticas sobre el programa de operación anual	Cantidad de biomasa cosechada en cada periodo Cantidad de biomasa procesada en cada periodo en cada refinera Cantidad de biocombustible producido en cada periodo en cada refinera Niveles de inventario de biomasa y residuos en cada periodo Flujo de Biomasa y residuos en la cadena
Eksioglu y Petrolia	2010	Análisis del impacto de una instalación intermodal en las decisiones de localización y transporte para las plantas de producción de biocombustible, en este caso Etanol proveniente de maíz.	Localización y capacidad de las biorrefinerías Localización y capacidad de los sitios de recolección Cantidad de biomasa cosechada en cada sitio, en cada periodo de tiempo Cantidad de biomasa y combustible almacenado en cada facilidad, en cada periodo de tiempo Cantidad de flujo en la red, en cada periodo de tiempo Modo de transporte utilizado Cantidad de biocombustible producido en cada biorrefinería en cada unidad de tiempo
Flores et. al.	2010	El modelo permite el análisis de la cadena productiva de biocombustibles, teniendo en cuenta el cultivo de las materias primas y la refinación de los biocombustibles.	Capacidades Rentabilidad Inversión Demandas Incentivos Precios
Huang et. al.	2010	El modelo matemático desarrolla la planeación estratégica, en un horizonte de tiempo, para la producción de biocombustibles a partir de diferentes biorresiduos.	Cantidad de flujo en la red en cada periodo Capacidad asignada a las biorrefinerías Asignación de biorrefinerías en cada periodo
Zapata et. al.	2010	El modelo estudia el comportamiento del aceite de palma y se propone diferentes políticas que ayuden a que se produzca un incremento en la oferta de biodiesel sin afectar lo que se dedica para el consumo humano.	Hectáreas sembradas de palma Hectáreas disponibles de palma Precios Costos Incentivos del gobierno Demandas Demandas no atendidas Inversiones Importaciones y exportaciones de palma

Eksioglu et. al.	2009	Determina el número, tamaño y ubicación de los bio-refinerías necesarias. Determina la cantidad de biomasa enviada, procesada e inventariada cada periodo de tiempo (Multi periodo)	Número de biorrefinerías Ubicación de las biorrefinerías Capacidad de las biorrefinerías Flujo de la biomasa en la red Niveles de inventario
Constantino, et. al	2008	Solución del modelo con Branch and bound (bifurcación y acotamiento) Restricciones de área máxima	Flujo en la agrocadena Asignación de actividades a los actores de la cadena
Lejars et. al.	2008	Planificación de la cosecha de caña de azúcar para mejorar la productividad y la ganancia	Cierre y reubicación de zonas de entrega Políticas de manejo del flujo de la caña en la cadena Técnicas utilizadas de cosecha y procesamiento
Ravula et.al.	2008	Secuenciación de las operaciones de transporte usando un número limitado de equipos especializados para la cosecha de la biomasa (algodón)	Asignación de Cosecha de cada región, por cada vía, en cada día de cosecha
Dunnett et. al.	2007	Determinar la selección de las facilidades logísticas, la capacidad y la asignación de tareas a las facilidades.	Utilización (o no) de cada facilidad logística en cada fase (momento) Asignación de tareas a cada facilidad logística Capacidad de cada refinería asignada en cada momento de tiempo Producción de cada refinería en cada momento de tiempo
Kumar y Sokhansanj	2007	Costos, energía de entrada y emisiones de carbono	Localización de cada tipo de planta Capacidad de cada tipo de planta Tasa de flujo de materia prima, producto intermedio y producto final en la red Cantidad de producto intermedio consumido en cada planta
Sokhansanj et. al.	2006	Simular la recolección, almacenamiento y operaciones de transporte para el abastecimiento de biomasa agrícola a la biorrefinería	Estado (húmedo o seca) y forma de recolección de la materia prima. Estado (húmedo o seca) y forma en que se transporta de la materia prima.
Martins et. al.	2005	Modelo de programación entera para un problema de cosecha forestal no-temporal con restricciones en el tamaño de la tala rasa y en el área total de parcelas de edad madura con un tamaño mínimo requerido.	Estructura de la zona de suministro del ingenio Duración de las temporadas de procesamiento Capacidad de la cadena por unidad de tiempo Reglas de asignación de despachos
Troncoso y Garrido	2005	Localización y tamaños óptimos de las facilidades logísticas forestales. Adicionalmente los niveles de producción y flujos a generar en un horizonte de planeación.	Posibilidad de cosecha de cada sección de cada tipo de bosque



Gunnarsson et. al.	2004	Cuándo y dónde los residuos forestales deben ser convertidos en combustible y cómo deben ser transportados y almacenados	Volumen de flujo en la red de cada presentación del producto Lugar de procesamiento de los productos
Higgins y Postma	2004	Incrementar la utilización de la infraestructura de transporte a través de la mejora en la secuenciación de las cosechas en la ferrovía y la carretera, suministrando la posibilidad de reducir los costos de transporte y cosecha.	Momento en que los módulos (dispositivos de proceso primario) son recogidos del campo de cultivo a la planta de proceso Días de programación del camión que recoge los módulos
Gigler et. al.	2002	Incluye dos características de los productos: Presentación (influenciada por la manipulación) y Calidad (influenciada por procesamiento, transporte y almacenamiento)	Asignación de recursos (mano de obra, equipo y estructura) Cantidad de biomasa producida y transportada
De Mol et. al.	1997	Desarrollo de dos modelos para obtener información sobre los costos y el consumo energético de la logística: uno de simulación y uno de optimización. La cadena estudiada incluye facilidades como: cultivos, recolección, transbordo, pre-tratamiento y plantas de energía.	Cantidad de hectáreas que se cortan por periodo en cada uno de las áreas de cultivo disponibles. Flujo en la red por periodo Ubicación de las instalaciones capacidad de las instalaciones Número de camiones utilizados por periodo Capacidad de expansión adicional en cada periodo
Cundiff	1996	Considera los problemas de almacenamiento, programación de cosechas y transporte	Flujo de biomasa en la cadena en cada momento de tiempo Capacidad de cada tipo de almacenamiento (almacén cubierto y a la intemperie) Cantidad de exceso y faltante de biomasa
Gallis	1996	Varios escenarios observados se utilizaron para definir el efecto del tiempo de inventario, la tasa de interés, y la pérdida de valor debido al deterioro de la fibra y los sistemas operativos sobre el costo por unidad y el costo total	Tiempo de inventario Tasa de interés Estado de la fibra Costo unitario y total

5.2 Cronología de los modelos matemáticos para la gestión de la cadena productiva del biodiésel

A partir de la revisión de la literatura sobre modelos matemáticos para la gestión de la cadena de biodiesel y cadenas afines que este documento ha considerado, se presenta una taxonomía de los mismos, basándose en la clasificación dada por CEPAL [11]. En la Figura 13, las líneas punteadas representan estudios aplicados a una cadena productiva específica, por lo que va referida a un número, que corresponde al número del artículo (ver Tabla V). Las líneas continuas representan la clasificación genérica de los tipos de biomasa y sus usos.

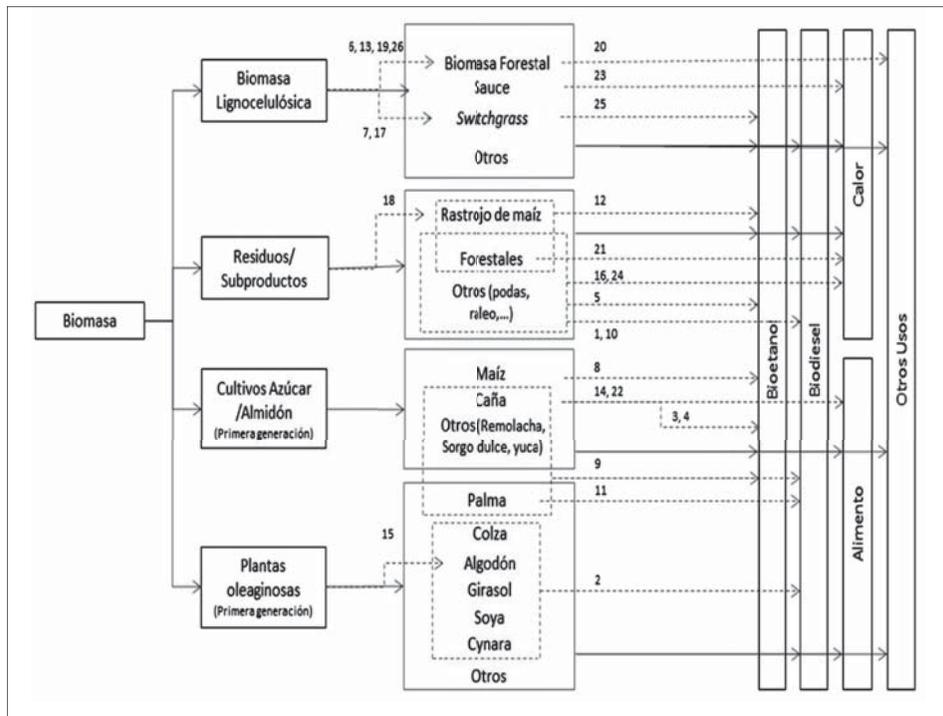


Figura 13. DTaxonomía de los modelos revisados.

Fuente: Los autores

En la Tabla X, se presentan los autores, las herramientas y el tipo de biomasa trabaja en los modelos estudiados en el presente estudio.

5.2.1. Biomasa Lignocelulósica

En este grupo se muestran nueve artículos, de los cuales cinco tratan de biomasa forestal, entendiéndose ésta como biomasa de bosque. Los otros artículos, se refieren a cadenas de Sauce y *Switchgrass* (conocido también como pasto varilla). A su vez, la mayoría de estos modelos se enfocan en la cadena productiva de la biomasa, sin especificar el uso final; sólo tres se dirigen a usos concretos, como obtención de bioetanol, generación de calor y la industria de la madera (clasificados como otros usos).

- Biomasa Forestal

En 1996 Gallis [31], en Grecia, realizó un modelo de simulación de la cadena de biomasa forestal de Grecia, que constituye uno de los recursos renovables existentes de las materias primas no sólo para la generación de energía, sino también para la industria de transformación de la madera. El objetivo fue evaluar los efectos de variables logísticas sobre el costo de la cadena mediante la observación de varios escenarios.



En 2005, Troncoso y Garrido [32], en Chile, proponen un modelo MILP de producción y logística de la industria forestal, donde se consideran variables de decisión sobre: el flujo de la cadena; el diseño de la cadena, como tamaño y ubicación de las facilidades logísticas; y de producción, como el momento de cosecha de cada sección (parcela) de cada tipo de terreno.

De manera independiente, en 2005 en Portugal, Martins et al. [33], proponen un modelo MILP para la programación de cosechas, con el fin de maximización del volumen de madera cortada. A diferencia de los anteriores estudios realizados, este modelo se centra en la parte operativa de la cadena productiva de la madera. Para resolver el mismo problema de programación de cosechas de madera, en 2008, Constantino et al. [34] presentaron un nuevo modelo que comprende un número polinomial de variables y restricciones, utilizando la técnica “bifurcación y acotamiento” para solucionarlo. Como último estudio realizado en cadenas forestales se encuentra el de Zhang et al. [35] en 2011 en el cual presenta un modelo MILP para el diseño y gestión de la cadena maximizando su Valor Presente Neto, disminuyendo el impacto causado al hábitat forestal y asegurando la sostenibilidad de los flujos de madera.

- Otra biomasa lignocelulósica

Cundiff [36], en 1996, plantea su modelo de programación lineal para la minimización del costo de la cadena de biomasa herbácea, aplicándolo en la cadena de *switchgrass*, considerando los costos de transporte, almacenamiento, y los costos de incertidumbre (faltante y exceso) de la producción causada por la variación del clima.

En 2007, Kumar y Sokhansanj [37] describen un modelo IB-SAL (*Integrated Biomass Supply Analysis and Logistics*) para cuantificar variables ecológicas como energía utilizada y huella de carbono en la cadena de biomasa de *switchgrass*.

Zhu et al. [38], en 2011, proponen un modelo matemático para la gestión de la cadena de biomasa a la biorrefinería, para la producción de biocombustible, en el cual se integran decisiones estratégicas para la planeación (a un año) y decisiones tácticas de programación de cosechas. Incluye características propias de la cadena de *switchgrass*: frecuencia y temporada de cosecha, variación del contenido de la biomasa según las condiciones del tiempo, efectos del clima y la distribución dispersa de los cultivos.

En 2002, Gigler et al. [39] desarrollaron un modelo de programación dinámica (PD) para la optimización de costos de agrocadenas, aplicándolo a biomasa de Sauce. Este modelo se basa en dos características propias de los productos agrícolas: Presentación en cada etapa de la cadena (árbol, tronco, trozos o astillas) y Calidad del producto, influenciada por las actividades de procesamiento, transporte y almacenamiento de la biomasa.

5.2.2. Cultivos de azúcar o almidón

En este grupo se enuncian cuatro artículos, de los cuales dos tratan la cadena productiva de la caña de azúcar para la producción de azúcar (alimentos). El otro se refiere a la cadena productiva de maíz para la producción de bioetanol.

En 2004, Higgins y Postma [40], de Australia, presentaron un modelo de MILP para la programación de cosechas, procesamiento primario y transporte de caña de azúcar, con el objetivo de maximización la rentabilidad, haciendo el mejor uso de sus recursos.

Lejars et al. [41], de Francia, en 2008 utilizaron simulación para planificar la cosecha de caña de azúcar con el fin de mejorar el rendimiento de esta, y por lo tanto, aumentar la ganancia de la cadena.

Posteriormente, en 2010, Eksioglu y Petrolia [42], desarrollaron otro modelo para el diseño y coordinación de la cadena de biomasa para la producción de biocombustibles, pero esta vez analizaron el impacto de diferentes modos de transporte en la cadena. El modelo se aplicó para la producción de bioetanol a partir de maíz.

Igualmente en 2010, Flórez et al. [43] exponen un modelo de dinámica de sistemas para analizar las políticas actuales de producción de biocombustibles y los cambios posibles para alcanzar la producción esperada. En este modelo se trabaja la producción de biocombustibles, siendo fuente de etanol la caña y la yuca.

En 2011, Orjuela et al. [3] presentan una caracterización de la caña panelera en Colombia, así como los proyectos vigentes y el desarrollo para la producción de biocombustibles en este país. Los datos históricos son tratados con estadística multivariada y redes neuronales, con el fin de predecir el comportamiento futuro de la producción de caña panelera. También se presenta un diagrama causal donde se identifican los factores y variables relevantes en este sistema productivo.

En el mismo año, Orjuela et al. [44] mediante dinámica de sistemas realizan un análisis de factores y variables relevantes en lo que respecta a la problemática sobre la repercusión que tienen los cultivos de caña panelera para bioetanol sobre la seguridad alimentaria de la población colombiana.

5.2.3. Plantas oleaginosas

En este grupo se muestran dos artículos. Uno hace referencia a la cadena de biomasa de algodón, sin especificar su uso final; el otro trata la cadena productiva de biodisel de primera generación a partir de diferentes materias primas.

En 2008, Ravula et al. [45] de Estados Unidos, desarrollaron un modelo de simulación discreta de la cadena de biomasa de algodón, con el fin de disminuir el tiempo de reco-



lección (desde los cultivos hacia las instalaciones de procesamiento) de la biomasa con restricciones de equipos especializados para la cosecha.

Flórez et al. [43] en 2010 en su modelo de dinámica de sistemas sobre producción de biocombustibles en Colombia, (explicado en la sección 5.2.2 de cultivos de azúcar/almidón) también consideran la producción de biodiesel a partir de palma de aceite. Igualmente, Zapata et al. [46] hacen un análisis en dinámica de sistemas específico sobre la producción de biodiesel en Colombia a partir de palma de aceite. Éste modelo pretende probar políticas nacionales que contribuyan al aumento de la oferta de biodiesel buscando no afectar lo que se dedica para el consumo humano.

Kim et al. [47], en 2011, plantearon un modelo MILP para el diseño y gestión de la cadena de biocombustibles, esta vez aplicado a la producción de biodiesel a partir de fuentes como residuos forestales, raleo, podas, hierbas y virutas. A diferencia del modelo de Papapostolou et al. [13], este modelo se enfocó en biomasa lignocelulósica, la cual requiere un proceso diferente del biodiesel de primera generación.

5.2.4. Subproductos y residuos

En este grupo se encuentran ocho artículos, de los cuales seis tratan la cadena productiva de biomasa de residuos; uno la de subproductos (rastreo de maíz); y el otro, trata conjuntamente la de residuos y subproductos. A su vez, la mayoría de estos modelos dirigen la biomasa a un uso específico: en esta caso para la producción de biodiesel, bioetanol o para generación de calor.

Sokhansanj et al. [48] en 2006, describen un modelo IBSAL para cuantificar los recursos necesarios y el costo de operación en el abastecimiento y transporte de la biomasa. Este modelo fue aplicado a la biomasa del rastreo de maíz.

En 2009, Eksioglu et al. [49], proponen un modelo matemático para la cadena de biomasa dirigida a biorrefinerías para la producción de biocombustibles. En este modelo se toman en cuenta variables de decisión a largo plazo (diseño de la cadena) y a mediano y corto plazo (gestión de la cadena), integrando estos dos niveles de decisión. El modelo es aplicado para la obtención de bioetanol a partir de rastreo de maíz y biomasa leñosa (residuos forestales, pulpa de madera y madera de aserrío).

En 1997 De Mol et al. [50], desarrollaron dos modelos: uno de simulación, para cuantificar los costos y consumo de energía en la cadena, y otro de optimización MILP, para Minimización el costo de la cadena. El objetivo de los modelos era el diseño de cadenas de biomasa de diferentes residuos de papel, madera, podas, raleo y sanitario para plantas de energía.

Gunnarsson et al. [51] en 2004, en Suecia, desarrollaron un modelo de programación lineal entera mixta (MILP por sus siglas en inglés) donde también trabajo la cadena de

biomasa forestal para obtener combustible forestal, pero esta vez la fuente fue residuos forestales. El objetivo de este modelo fue minimización el costo total generando una estrategia de distribución y procesamiento de los residuos.

Dunnet et al. [52] en 2007, presentaron un modelo MILP enfocado al diseño de la cadena de biomasa para centrales de combustión en la que se asignan las facilidades logísticas a utilizar en cada momento de tiempo, especificando sus capacidades y tareas asignadas: recolección, compresión, secado, almacenamiento y transporte de la biomasa.

Por su parte, Huang et al. [29] en 2010 formularon un modelo MILP para la planeación de operaciones en un horizonte de tiempo determinado, para la producción de biocombustibles a partir de biorresiduos. Este modelo se aplicó para evaluar el potencial económico y las necesidades de infraestructura para la producción de bioetanol a partir de residuos de biomasa en California, USA.

Así mismo, Papapostolou et al. [13], en Grecia, en 2011, propusieron un modelo MILP genérico aplicado al diseño y gestión de la cadena de biocombustibles, considerando opciones de producción o importación de materias primas. El modelo se implementó en Grecia, en la cadena de biodiesel de primera generación, proveniente de materias primas como colza, algodón, girasol, cynara y soya.

Por último Chien y Fan [53], en 2011, plantearon un modelo de programación estocástica entera mixta para soportar la planeación estratégica de cadenas productivas de bioetanol a partir de residuos. Este modelo involucra tanto el diseño de la cadena (localización y capacidad de refinerías y terminales) y coordinación de la cadena (plan de transporte) en un ambiente de decisión, bajo incertidumbre.

6. Clasificación de los modelos matemáticos para SCM de petrocombustibles

An et al. [30] en 2011 evalúan artículos referentes a modelos matemáticos aplicados a cadenas productivas de combustible basado en petróleo. Al igual que con los modelos de biocombustibles, los artículos son categorizados de acuerdo a dos dimensiones (nivel en la cadena de suministro y nivel de decisión) como se muestra en la Figura 14.

Los modelos matemáticos desarrollados para la toma de decisiones en las cadenas productivas de combustible basado en petróleo se enfocan principalmente con los niveles medio y bajo de la cadena, es decir, “*midstream*” y “*downstream*”, correspondientes a los procesos de refinado y distribución a las estaciones de servicio. Adicionalmente, se evidencia una mayor concentración en modelos que apoyan las decisiones a nivel táctico; sin embargo, también se presentan modelos para los niveles operativo e integrado y ninguno para el nivel estratégico.

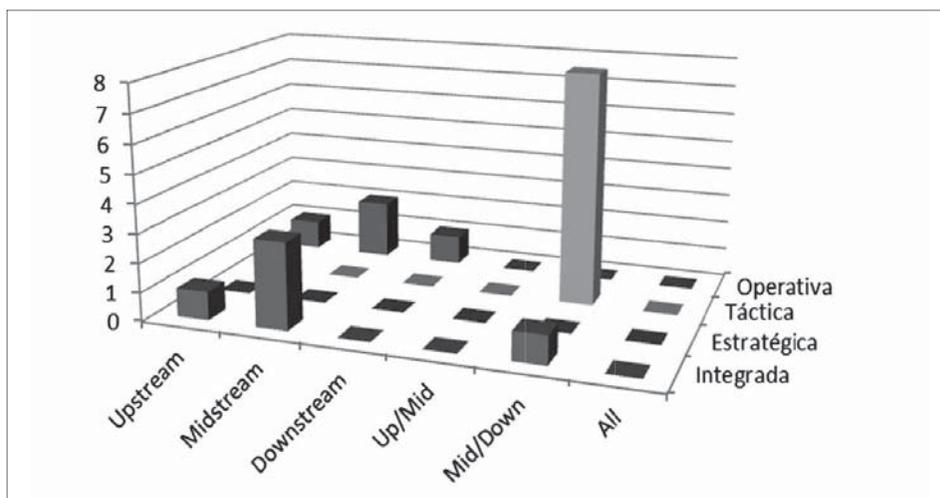


Figura 14. Categorización de los artículos relacionados a SCM de combustibles basados en petróleo.

Fuente: A partir de [30]

Conclusiones

En la revisión de la literatura no se encontraron artículos científicos de modelos matemáticos de SCM de biocombustibles. La mayoría de artículos provienen de países refinadores de biocombustibles, como USA y varios países de la Unión Europea. Pero prácticamente nada de los países más productores de biomasa son Malasia, Indonesia (Oriente), México, Brasil y Argentina (América). Así mismo, no se encontró ningún artículo para el diseño de la cadena productiva de biodiesel a partir de palma de aceite. En razón de lo anterior, se presenta una referencia importante en este trabajo.

Es evidente que la herramienta más utilizada en el diseño de modelos matemáticos de SCM es la programación lineal entera mixta (MILP). La mayor parte de los modelos buscan apoyar la toma de decisiones en el diseño de la cadena (localización y plan de ampliación de capacidades), en las cuales se utilizan variables de decisión booleanas (variables enteras); y en la gestión de la cadena que contemplan los flujos a largo de la cadena desde materias primas, productos intermedios y productos terminados (como variables continuas).

Los modelos matemáticos para la cadena de los biocombustibles son diferentes a los de la cadena de los combustibles basados en petróleo. Mientras los modelos matemáticos para el biocombustible se concentran en un nivel alto de la cadena (*upstream*), los combustibles basados en petróleo se concentra en los niveles medio y bajo de la cadena (*downstream* y *midstream*). Esto probablemente debido a que mientras en la producción del biocombustible son importantes las decisiones relacionadas con la materia prima (a nivel operativo, táctico y estratégico), en el petróleo las decisiones en la cadena de los petro-

combustibles se dirigen a los procesos de refinamiento y distribución, y a la extracción en yacimientos.

A partir de la revisión realizada en este estudio, se evidencia que la mayoría de modelos matemáticos se concentran en el diseño y gestión de cadenas de biomasa de residuos y plantas lignocelulósicas. Pocos modelos tratan cadenas de biocombustibles de primera generación (plantas oleaginosas para la producción de biodiesel y de azúcar o almidón para la producción de etanol). Lo anterior sugiere un campo de investigación abierto en relación a desarrollar modelos matemáticos de SCM para el biodiesel, que sirvan para apoyar la toma de decisiones en el diseño y gestión de estas cadenas productivas.

Referencias bibliográficas

- [1] DUFÉY, Annie. (2006). Producción y comercio de biocombustibles y desarrollo sustentable: los grandes temas. Instituto Internacional para el Medio Ambiente y Desarrollo. Londres. Documento de Discusión Número 2 de Mercados Sustentables.
- [2] ARÍSTEGUI, Juan Pablo. (2009). Los biocombustibles desde la perspectiva del comercio internacional y del derecho de la organización mundial del comercio. *Revista de Derecho (Austral)*, vol. XXII, núm. 1, julio, 2009, pp. 113-134 Universidad Austral de Chile.
- [3] ORJUÉLA, Javier. HUERTAS, Isaac. FIGUEROA, Juan. KALENATIC, Dusko. CADENA, Katherine. (2011). Potencial de producción de bioethanol a partir de caña panelera: Dinámica enre contaminación, seguridad alimentaria y uso del suelo. Universidad Distrital Francisco José de Caldas. *Revista de Ingeniería Vol 16, N° 1*. Bogotá.
- [4] BALAT, Mustafa. (2010). Potential alternatives to edible oils for biodiesel production – A review of current work. Trabzon, Turquía. *Energy Conversion and Management* 52 1479–1492.
- [5] IICA. (2010). Atlas de la Agroenergía y los Biocombustibles en las Américas: II Biodiesel. - Instituto Interamericano de Cooperación para la Agricultura. Costa Rica.
- [6] KRAWCZYK, T. (1996). Biodiesel - alternative fuel makes inroads but hurdles remain. *INFORM*;7:800–15
- [7] BENJUMEA, Pedro. AGUDELO, John. RÍOS, Luis. (2009). Biodiésel: Producción, calidad y caracterización. Editorial Universidad de Antioquia. Medellín, Colombia.
- [8] NG, J.-H. NG, Hoon Kiat, GAN, S. (2009). Recent trends in policies, socioeconomy and future directions of the biodiesel industry. Malasia. *Clean Techn Environ Policy* 12: 213–238
- [9] DEMIRBAS, Ayhan. (2009). Progress and recent trends in biodiesel fuels. Trabzon, Turquía. *Energy Conversion and Management* 50 14–34.
- [10] USDA. (2012). Oilseeds: World Markets and Trade. United States Department of Agriculture. Foreign Agricultural Service. Circular Series FOP 06-12.
- [11] CEPAL. (2011). Análisis comparativo de patentes en la cadena de producción de biocombustibles entre América Latina y el resto del mundo. Santiago, Chile. Comisión económica para América Latina y el Caribe. Diálogo de Políticas sobre desarrollo institucional e innovación en biocombustibles en América Latina y el Caribe Santiago de Chile, 28 y 29 de marzo, 2011.
- [12] ORJUÉLA, Javier. RAMÍREZ, Carmen. LINARES, Karina. (2010). Procesos productivos para la obtención de agro combustibles. Universidad Católica de Colombia. Studiositas, edición de diciembre, vol5, núm.3.



- [13] PAPAPOSTOLOU, Christiana. KONDILI, Emilia. KALDELLIS, John K. (2011). Development and implementation of an optimisation model for biofuels supply chain. Grecia. Energy 36 6019 - 6026.
- [14] MITTELBAACH, M. REMSCHMIDT, C. (2004). Biodiesels – the comprehensive handbook. Graz, Austria: Karl-Franzens University Press.
- [15] KNOTHE, G. SHARP, CA. RYAN, TW. (2006). Exhaust emissions of biodiesel, petrodiesel, neat methyl esters, and alkanes in a new technology engine. Energy Fuels; 20:403–8.
- [16] DEMIRBAS, Ayhan. (2007). Importance of biodiesel as transportation fuel. Trabzon, Turquía. Energy Policy 35 4661–4670.
- [17] DEMIRBAS, Ayse Hilal. DEMIRBAS, Imren. (2007). Importance of rural bioenergy for developing countries. Trabzon, Turquía. Energy Conversion and Management 48 2386–2398.
- [18] YEE, Kian Fei. TAN, Kok Tat, ABDULLAH. Ahmad Zuhairi. LEE, Keat Teong. (2009). Life cycle assessment of palm biodiesel: Revealing facts and benefits for sustainability. Pulau Pinang, Malasia. Applied Energy 86 S189–S196.
- [19] FAZAL, M.A. HASEEB, A.S.M.A. MASJUKI, H.H. (2010). Biodiesel feasibility study: An evaluation of material compatibility; performance; emission and engine durability. Kuala Lumpur, Malasia. Renewable and Sustainable Energy Reviews 15 1314–1324.
- [20] CHOUINARD, Pascale. BRADT, Laura. PONCE José M. EL-HALWAGI, Mahmoud M. (2010). Incorporation of process integration into life cycle analysis for the production of biofuels. USA y Mexico. Clean Technologies and Environmental Policy. Volume 13, Number 5, 673-685.
- [21] BALA, BK. (2005) Studies on biodiesels from transformation of vegetable oils for diesel engines. Energy Educ Sci Technol.15:1–43.
- [22] AJANOVIC, A. HAAS, R. (2010). Economic challenges for the future relevance of biofuels in transport in EU countries. Austria. Energy 35 3340 - 3348.
- [23] FEDEBIOCOMBUSTIBLES. (2012). Cifras Informativas del Sector Biocombustibles - biodiésel de palma de aceite. Federación nacional de biocombustibles de Colombia.
- [24] MINAGRICULTURA. (2009). Política Nacional de Biocombustibles - Visión desde el sector Agropecuario. Ministerio de Agricultura y Desarrollo Rural de la República de Colombia.
- [25] RODRÍGUEZ, Isaías. ORJUELA, Javier. (2009). Panorama de las políticas y leyes del gobierno colombiano frente a la producción de alimentos agropecuarios y de producción de agrocombustibles. Estudios en derecho y gobierno, Vol. 2, N° 2. Universidad Católica de Colombia.
- [26] GUALTEROS S, Juan Manuel. (2011). Estudio prospectivo de la cadena productiva del biodiésel a partir de palma africana en Colombia. Universidad Nacional. Bogotá, Colombia.
- [27] NACEPT. (2007). Strategic Framework for Biofuels Efforts. National Advisory Council for Environmental Policy and Technology. Washington DC, Estados Unidos.
- [28] AWUDU , Iddrisu. ZHANG, Jun. (2012). Uncertainties and sustainability concepts in biofuel supply chain management: A review. Estados Unidos. Renewable and Sustainable Energy Reviews 16 1359– 1368.
- [29] HUANG, Yongxi. CHEN, Chien-Wei. FAN, Yueyue. (2010). Multistage optimization of the supply chains of biofuels. Estados Unidos. Transportation Research Part E 46 820–830.
- [30] AN, Heungjo. WILHELM, Wilbert E. SEARCY, Stephen W. (2011). Biofuel and petroleum-based fuel supply chain research: A literature review. Estados Unidos. biomass and bioenergy 35 3763 – 3774.
- [31] GALLIS, Christos TH. (1996). Activity oriented stochastic computer simulation of forest biomass logistics in Greece. Forest Engineering Researcher, Greece. Biomass and Bioenergy Vol. 10. Nos 516, pp. 377-382.

- [32] TRONCOSO, Juan J. GARRIDO, Rodrigo A. (2005). Forestry production and logistics planning: an analysis using mixed-integer programming. Pontificia Universidad Católica de Chile, Santiago de Chile, Chile. *Forest Policy and Economics* 7: 625– 633.
- [33] MARTINS, Isabel. CONSTANTINO, Miguel. BORGES, José G. (2005). A column generation approach for solving a non-temporal forest harvest model with spatial structure constraints. Lisboa, Portugal. *European Journal of Operational Research* 161: 478–498.
- [34] CONSTANTINO, Miguel. MARTINS, Isabel. BORGES, José G. (2008). A New Mixed-Integer Programming Model for Harvest Scheduling Subject to Maximum Area Restrictions. Lisboa, Portugal. *OPERATIONS RESEARCH*, Vol. 56, No. 3, May–June, pp. 542–551.
- [35] ZHANG, Huizhen. CONSTANTINO, Miguel. FALCÃO, André. (2011). Modeling forest core area with integer programming. Centro de Investigação Operacional, Universidade de Lisboa, Portugal. *Ann Oper Res* 190:41–55.
- [36] CUNDIFF, John S. DIAS, Neil. SHERALI, Hanif D. (1996). A linear programming approach for designing a herbaceous biomass delivery system. Virginia Polytechnic Institute and State University, Blacksburg, VA, USA. *Bioresource Technology* 59: 47-55.
- [37] KUMAR, Amit. SOKHANSANJ Shahab. (2007). Switchgrass (*Panicum virgatum*, L.) delivery to a biorefinery using integrated biomass supply analysis and logistics (IBSAL) model. *Bioresource Technology* 98 (2007) 1033–1044.
- [38] ZHU, Xiaoyan. LI, Xueping. YAO, Qingzhu. CHEN, Yuerong. (2011). Challenges and models in supporting logistics system design for dedicated-biomass-based bioenergy industry. USA. *Bioresource Technology* 102. 1344–1351.
- [39] GIGLER, J.K. HENDRIX, E.M.T. HEESSEN, R.A. VAN DEN HAZELKAMP, V. MEERDINK, G. (2002). On optimisation of agri chains by dynamic programming. Wageningen, Netherlands. *European Journal of Operational Research* 139: 613–625.
- [40] HIGGINS, Andrew J. POSTMA, Steve. (2004). Australian Sugar Mills Optimise Siding Rosters to Increase Profitability. Australia. *Annals of Operations Research* 128, 235–249.
- [41] LEJARS, Caroline. LE GAL, Pierre-Yves. AUZOUX, Sandrine. (2008). A decision support approach for cane supply management within a sugar mill area. Francia. *computers and electronics in agriculture* 60: 239–249.
- [42] EKŞIOĞLU, Sandra D. PETROLIA, Daniel. (2010). Analyzing the Impact of Intermodal Facilities to the Design and Management of Biofuels Supply Chain. Mississippi State University, USA.
- [43] FLÓREZ B., Ana M. , FRANCO C., Carlos J. , IDYNER, saac . (2010). Análisis de la producción de biocombustibles en Colombia. Universidad Nacional de Colombia. 8° Congreso Latinoamericano y 8° Encuentro Colombiano de Dinámica de Sistemas.
- [44] ORJUELA C, Javier A., CADENA O., Katherine, BERNAL H., Milena, FIGUEROA G., Juan C. HUERTAS, Isaac . (2011). Dinámica de Producción de Bioetanol a partir de la Caña Panelera, Incidencia en la Seguridad Alimentaria. IX Congreso Latinoamericano de Dinámica de sistemas y II Congreso Brasileño de Dinámica de Sistemas.
- [45] RAVULA, Poorna P. GRISSO, Robert D, CUNDIFF John S. (2008). Cotton logistics as a model for a biomass transportation system. Biological Systems Engineering, Virginia Tech, USA. *BIOMASS AND BIOENERGY* 32: 314 – 325.
- [46] Sebastián ZAPATA R. Carlos J. FRANCO C. Ana M. FLÓREZ B. (2010). Modelo de producción de palma de aceite. 8° Congreso Latinoamericano y 8° Encuentro Colombiano de Dinámica de Sistemas.
- [47] KIM, Jinkyung. REALFF Matthew J. LEE, Jay H. WHITTAKER Craig. FURTNER, Ludwig. (2011). Design of biomass processing network for biofuel production using an MILP model. *Biomass and bioenergy* 35 853 – 871.



- [48] SOKHANSANJ Shahab, KUMAR Amit, TURHOLLOW, Anthony F. (2006). Development and implementation of integrated biomass supply analysis and logistics model (IBSAL). *Biomass and Bioenergy* 30: 838–847.
- [49] EKŞIOĞLU, Sandra D. ACHARYA, Ambarish. LEIGHTLEY, Liam E. ARORA, Sumesh. (2009). Analyzing the design and management of biomass to biorefinery supply chain. *USA. Computers & Industrial Engineering* 57: 1342–1352.
- [50] DE MOL, R.M. JOGEMS, M.A.H. VAN BEEK, P. GIGLER, J.K. (1997). Simulation and optimization of the logistics of biomass fuel collection. Wageningen, Netherlands. *Netherlands Journal of Agricultural Science* 45: 219-228.
- [51] GUNNARSSON, Helene. RÖNNQVIST, Mikael. LUNDGREN, Jan T. (2004). Supply chain modelling of forest fuel. Linköping Institute of Technology, Suecia. *European Journal of Operational Research* 158: 103–123.
- [52] DUNNETT, A. ADJIMAN, C. SHAH, N. (2007). Biomass to heat supply chains Applications of Process Optimization. Centre for Process Engineering, Department of Chemical Engineering, Imperial College, London, UK.
- [53] CHEN, Chien-Wei. FAN, Yueyue. (2011). Bioethanol supply chain system planning under supply and demand uncertainties. *Universidad de California. USA. Transportation Research Part E* 48 150–164.

Manuel José Barón Molina

Es Ingeniero Industrial de la Universidad Autónoma de Colombia, de Bogotá, Colombia. Cursó una especialización en Gerencia de Negocios Internacionales en la Universidad Jorge Tadeo Lozano, en Bogotá, Colombia. Actualmente realiza su investigación para obtener su título de Maestría en Ingeniería Industrial en la Universidad Católica de Colombia, en Bogotá, Colombia.

Actualmente se desempeña como profesor en las áreas de Logística e Investigación de Operaciones, en la Universidad Católica de Colombia. e-mail: mjbaron@ucatolica.edu.co

Isaac Huertas Forero

Es Estadístico Universidad Nacional de Colombia, Bogotá. Obtuvo su título de Maestría en Investigación y Estadística en la Universidad Tecnológica de Pereira, Colombia.

Actualmente se desempeña como profesor en las áreas de Logística e Investigación de Operaciones en la Universidad Católica de Colombia y Universidad Distrital Francisco José de Caldas. Pertenece como líder al grupo GEGI, en la Universidad Católica de Colombia, en Bogotá, Colombia. e-mail: ihuertas@ucatolica.edu.co

Javier Arturo Orjuela Castro

Es Ingeniero de Alimentos, Ingeniero Industrial, Especialista en Ingeniería de producción de la Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Obtuvo su título de Maestría en Investigación y Estadística en la Universidad Tecnológica de Pereira, Colombia.

Actualmente se desempeña como profesor en las áreas de Logística e Investigación de Operaciones en varias universidades a nivel de pregrado y posgrado en la Universidad Católica de Colombia y Universidad Distrital Francisco José de Caldas. e-mail: jaorjuela@etb.net.co

Una revisión al estado del arte de la integración de toma de decisiones en la red logística

State of the art review of the integration of decision making in the logistics network

Cesar Augusto Pineda Pérez

Magister en Ingeniería Industrial
Universidad Distrital Francisco José de Caldas
capinedap@udistrital.edu.co

Cesar Amílcar López Bello

Docente
Universidad Distrital Francisco José de Caldas
Facultad de Ingeniería
clopezb@udistrital.edu.co



Resumen

El presente artículo es resultado de la investigación del proyecto titulado "Integración de Toma de decisiones en la red logística" y comprende la revisión del estado del arte en el cual se identifica el concepto de integración en la cadena de suministro, los tipos de integración, mecanismos de integración, herramientas de integración, estrategias de integración y finalmente los beneficios y las medidas de rendimiento.

Palabras clave: Integración de la cadena de suministro, tecnologías de información, arcos de integración, outsourcing de servicios.

Abstract

This article is the result of the research project entitled "Integration of Decision Making in the logistics network". It conduct a review of the state of art in which the concept of integration in the supply chain is identified, as well as the types of integration, mechanisms of integration, integration tools, integration strategies and finally the benefits and performance measures.

Key words: Integration of the supply chain, information technologies, arcs of integration, outsourcing services.

1. Introducción

En el escenario de la manufactura global, las decisiones en la cadena de suministro no se pueden tomar por separado. Hoy en día estas decisiones tienen que ver con los procesos de integración dentro y fuera de la empresa. Proveedores, plantas de manufactura, centros de servicio,

Fecha recibido: may. 8/2013
Fecha modificado: jun. 10/2013
Fecha aceptado: jun. 14/2013



operadores logísticos, contratistas, distribuidores, minoristas, detallistas, en fin, todos los actores de una red logística o de redes logísticas tienen aportan a los procesos de integración y asimismo a tomar decisiones conjuntas que beneficien a la red y a todos los actores que la integran (Taskin Gumus, 2007), (Ahmed Musa, 2013).

Los procesos de integración de decisiones en la red logística se apoyan en gran medida con el avance de las tecnologías de la información, esto hace que la manufactura global esté dentro del alcance de muchas organizaciones pequeñas que entrarán a engrosar la alta competencia de mercado (Vernadat, 2002), (Brent D. Williams, 2013).

Una de las decisiones conjuntas de gran trascendencia es la minimización de los costos logísticos. Los costos logísticos dominan los gastos de las empresas manufactureras. Por lo tanto, las decisiones logísticas involucran la obtención de la materia prima, producción y distribución y estas decisiones no se deben hacer por separado sino integradas con el fin de obtener ahorros. Para maximizar los ahorros, varias investigaciones han propuesto la integración de las decisiones logísticas de una red producción-distribución dentro de un ambiente de manufactura global. Así, una empresa de manufactura global puede ser formulado como una combinación red producción-distribución constituida de múltiples fuentes y múltiples destinos (Pagell, 2004), (Christina W.Y. Wong, 2013).

En recientes estudios, los modelos para coordinar redes de sistemas de producción-distribución han tendido a focalizar la integración de decisiones en programación de transporte, políticas y control de inventarios y programación de tamaños de lotes de producción (Ahmed Musa, 2013), (Brent D. Williams, 2013).

El presente artículo se enfoca específicamente en describir el estado del arte en cuanto a la integración de la toma de decisiones en la cadena de suministro desde sus conceptos primarios hasta la justificación de la integración.

Metodología

Para abordar el estado del arte en cuanto a la Integración de la Toma de decisiones en la Red Logística se consultaron diferentes bases de datos bibliográficas (Scopus, Science-Direct, SpringerLink, Proquest, IEEE, IOPSCIENCE). De los artículos encontrados, se seleccionaron los artículos que tratan específicamente el tema. Para construir el estado del arte de la integración de toma de decisiones en la red logística, se requirió responder las siguientes preguntas: ¿Qué se entiende por integración?, ¿Qué se puede integrar?, ¿Qué tipos de integración existen?, ¿Cómo se integra?, ¿Qué se necesita integrar?, ¿Para qué se integra?, ¿Qué estrategias de integración existen?, ¿Cómo se mide la integración? Con base en estas preguntas se estructuró el estado del arte como se expone en la Sección 3.

3. Marco teórico

La integración de la cadena de suministro ha sido tratada desde diferentes perspectivas. Es así, como se distinguen la integración con clientes, integración de la información, integración logística y de distribución e integración con proveedores (Vijayasathy, 2010), (Christina W.Y. Wong, 2013), (Claudia Colicchia, 2013), (Danese, 2013).

3.1. ¿Qué se entiende por integración?

Para definir qué es integración, no existe una definición única y aceptada o puesta en funcionamiento. La literatura e investigación sobre el concepto de integración de la cadena de suministro es desarrollada por varios investigadores. Entre las definiciones más conocidas podemos citar las siguientes (Pagell, 2004), (Xiande Zhao, 2008), (Devendra Choudhary, 2013), (Edrisi Muñoz, 2013):

- Kahn y Mentzer (1998) y Pinedo (2005) definen la integración como el proceso de integración entre los departamentos y la colaboración interdepartamental que aporta a una organización cohesiva.
- O’Leary Kelly y Flores (2002) afirman que la integración se refiere al grado en el cual las partes trabajan juntas de manera cooperativa para llegar a resultados mutuamente aceptables.
- O’Leary Kelly y Flores (2002) y Kahn y Mentzer (1998) a partir de un esfuerzo de investigación definieron la integración así: “la integración es un proceso de interacción y colaboración en el que la fabricación, compras y logística trabajan juntos de manera cooperativa para llegar a resultados mutuamente aceptables para su organización”.
- Según Bowersox y Muran (1989), Hillerbrad y Braman (2003), definen la integración de la cadena de suministro como el grado en que el fabricante colabora en forma estratégica con sus socios de la cadena de suministro y gestiona los procesos de colaboración intra e inter-organizacional.
- Stank et.al (2001), enuncia que la integración de la cadena de suministro consiste en la integración interna de diferentes funciones dentro de una compañía e integración externa con los socios comerciales.

Si bien el concepto básico de la integración puede ser familiar para la mayoría de los investigadores, las operacionalizaciones reales de la construcción del concepto de integración serían muchas. La siguiente Tabla I, resume los estudios representativos relacionados con la integración y las operacionalizaciones de la construcción utilizadas en estos estudios (J.F., 1981), (Chien, 1993), (Speranza, 1994), (Hall, 1996), (Thomas, 1996), (Fumero, 1997), (Bookbinder, 1999), (Pagell, 2004), (Sanchez, 2006), (Farahani, 2013).



Tabla I. Taxonomía - Usos representativos de la construcción del concepto de la integración de la cadena de suministro

Referencia	Enfoque del estudio	Nivel de análisis	Definición
Ettlie y Stoll (1990)	Desarrollo de productos	Interno	Coordinación y colaboración
Alder (1992)	Desarrollo de productos	Interno	Coordinación
Sussman y Dean (1992)	Desarrollo de productos	Interno	Dos funciones de trabajo en conjunto
Wheelwright y Clark (2005)	Desarrollo de productos	Interno	Integración entre los ingenieros de diseño con los ingenieros de procesos para resolver problemas en el desarrollo de productos
Kahn y McDonough (1997)	Manufactura y mercadeo	Interno	Dos componentes que se integran e interactúan
Mintzberg et.al. (1996)	Todas las formas de colaboración	Interno y externo	Colaboración = Integración
Kahn y Mentzer (1998)	Mercadeo y otros departamentos	Interno	Interacción y colaboración (Kahn y McDonough, 1997). La integración se define formalmente como un proceso de interacción y colaboración interdepartamental.
Ellinger et.al. (2000)	Mercadeo y logística	Interno	Kahn y McDonough (1997) -dos grandes dimensiones- colaboración e interacción
Frolich y Westbrook (2001)	Cadena de suministro	Externo	Coordinación
Krajewskis y Wei (2001)	Proveedor comprador	Externo	Colaboración y coordinación
Narasingham y Das (2001)	Compras en la empresa	Interno	La integración de compras implica la participación activa de la empresa y tiene como objetivo promover la alineación de las metas y prácticas de compras con las prioridades estratégicas de negocio
Verma et.al. (2001)	Operaciones y mercadeo	Interno	El enfoque de trabajar en forma conjunta las operaciones y el marketing mejoraran sus propias funciones y relaciones.
Ganeshan (2002)	Papel de la tecnología Internet en la cadena de suministro	Cadena de suministro (interno y externo)	Plataforma tecnológica compartida
Gerwin y Barrowman (2002)	Desarrollo de productos	Interno	IPD es un enfoque de gestión para mejorar el rendimiento del desarrollo de nuevos productos
Kelly (2002)	ERP	Cadena de suministro (Interno y externo)	Plataforma tecnológica compartida

Naransinham y Kim (2001)	Cadena de suministro	Interno y externo	Coordinación
O'Leary-Kelly y Flores (2002)	Manufactura/Mercados	Interno	La integración se refiere a la medida en que las partes separadas trabajan juntos en forma cooperativa para llegar a resultados aceptables para ambas partes. En consecuencia, esta definición abarca construcciones relacionadas con el grado de cooperación, coordinación, interacción y colaboración
Swan (2002)	Desarrollo de productos	Interno	Coordinación

Existe un amplio cuerpo de investigación sobre las relaciones unidimensionales de la Integración de la cadena de suministro, examinando las relaciones de colaboración entre el fabricante y/o clientes o proveedores (Paulra, et.al 2008; Mabert y Ventiaramanan, 1998; Spekman et.al. 1998; Fawcett y Magnan, 2002). Mientras que algunos se centran en las relaciones diádicas con socios de la cadena de suministro (Lee y Whang, 2001), otros se centran en la gestión de la cadena de suministro como un sistema único, en lugar de intentar optimizar individualmente los subsistemas fragmentados (Vickery et.al. 2003; Naylor et.al. 1999; Bowersox y Murach, 1989; Hammer, 1990; Stevens, 1989). Si bien algunas definiciones de la integración de la cadena de suministro (SCI) hacen énfasis en los flujos de materiales y piezas, otros se centran en los flujos de información, recursos y dinero en efectivo. Se construye sobre la literatura existente de la integración de la cadena de suministro, incluido el fabricante (integración interna) y se extiende en ambas direcciones (integración de clientes y proveedores) (Blumenfield, 1985), (Benjamin, 1989), (Sharp, 1989), (Cohen, 1988), (Blumenfield, 1991), (Neal Juster, 1997), (Wang Sena, 2004), (Pagell, 2004), (Ferdinand Jaspers, 2006), (Farahani, 2013).

3.2. ¿Qué se puede integrar?

Para responder a la pregunta ¿Qué se puede integrar?, evidentemente se relaciona con la pregunta ¿Qué se puede coordinar en la cadena de suministro o red logística? Los artículos que abordan la planeación coordinada entre dos o más etapas de la SC (Supply Chain) se clasifican según la Figura 1 (Arshinder, 2008).

De acuerdo al tipo de coordinación se relacionan los ítems (inventarios, producción, lotes de producción, pedidos, transporte, distribución, almacenamiento, etc) que se pueden integrar. Algunos de estos tipos de coordinación se describen a continuación (Mithun Sharma, 2008), (Arshinder, 2008):

- **Coordinación entre Producción y Distribución.** La integración de estas dos funciones puede dar lugar a un ahorro sustancial en los costos globales y una mejora en el servicio mediante la explotación de economías de escala de producción y transporte. Algunos de los ítems que se pueden integrar son: El lote de producción, Cargas de vehículos, Control y manejo de los inventarios, Rutas de

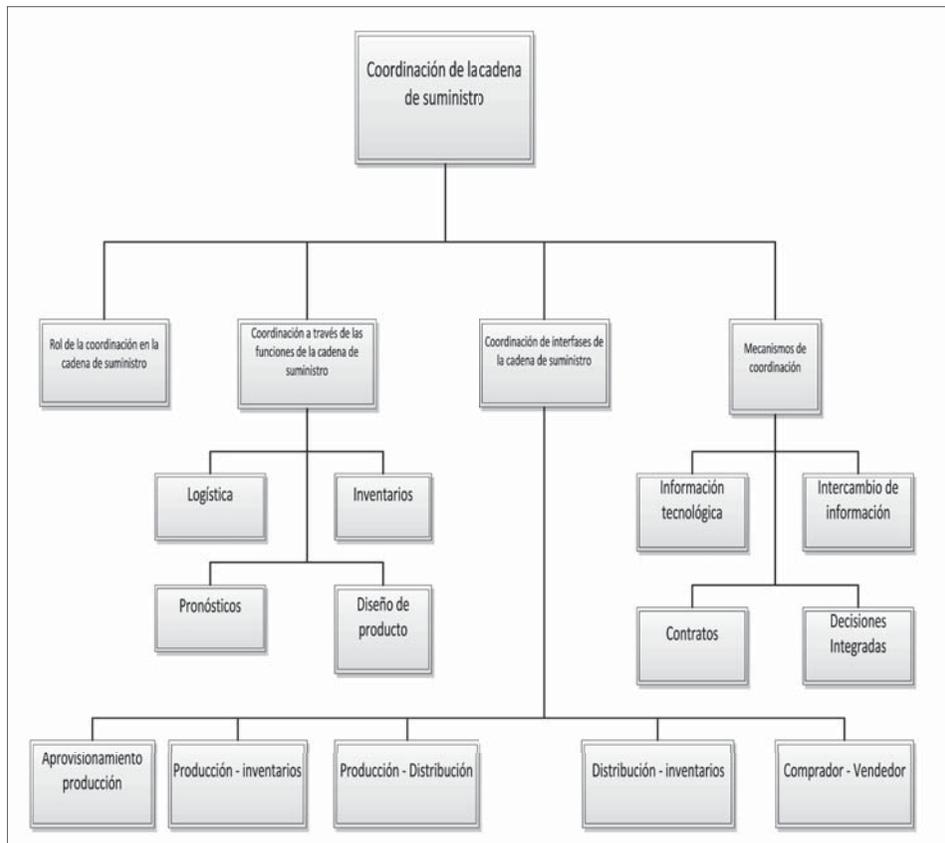


Figura 1. Taxonomía Tipos de Coordinación de la cadena de suministro. Fuente: (Arshinder, 2008)

distribución, Longitud del ciclo de producción, Frecuencia de pedidos, Tamaño del pedido, Asignación de producción.

- Coordinación entre aprovisionamiento y producción. Goyal y Deshmukh (1992) revisaron la literatura sobre los sistemas integrados de producción y aprovisionamiento (IPP). Los diferentes modelos IPP se clasificaron en categorías basadas en el número de productos, horizonte de planeación, métodos de solución empleados, ordenes de reposición y soluciones algorítmicas. Algunos de los ítems que se pueden integrar son: Control y manejo de inventarios, Frecuencia de pedidos de materia de prima o partes de ensamble, Tamaño de los pedidos de la materia prima o partes de ensamble.
- Coordinación entre producción e inventario. Lu (1995) consideró el enfoque heurístico del problema para un único vendedor con múltiples compradores con envíos de igual tamaño. Con la coordinación de las reposiciones de los distintos elementos, el vendedor puede reducir su costo total anual en un 30%. Los compradores también se benefician del modelo multi-comprador, redu-

ciendo sus costos. Algunos de los ítems que se pueden integrar son: Reposiciones, Políticas de inventario, Niveles de reorden, Stock de seguridad, Tiempos de suministro.

- Coordinación entre distribución e inventario. Jayaraman (1998) desarrollaron un modelo integrado de programación matemática mixta-entera para la minimización del costo total de distribución asociado con tres componentes de decisión: ubicación de las instalaciones, los parámetros de inventarios y alternativas de selección de transporte. Yokoyama (2002) desarrolló un modelo de optimización integrada del sistema producción-inventario en la que cualquier punto de consumo puede ser suministrado por los centros de distribución múltiples. Las políticas de inventario de revisión periódica y el problema de transporte se consideran al mismo tiempo. Algunos de los ítems que se pueden integrar son: Ubicación de instalaciones (centros de acopio, almacenes, bodegas, centros de distribución), Parámetros de inventarios, Modos de transporte, Selección de rutas, Cantidades a transportar, Número de envíos, Cantidad a ordenar.

Diversas perspectivas se presentan en la literatura sobre la coordinación de la cadena de suministro. La Tabla II resume varias perspectivas de la coordinación de la cadena suministro (Arshinder, 2008).

Los problemas de coordinación en las interfaces de la cadena de suministro se presentan en la Tabla III. Sharufali (2000) hizo énfasis en las ventajas de los mecanismos de coordinación tales como cambios en los precios, descuentos por cantidad (Sarmath et.al 2007), entregas parciales y el establecimiento de políticas conjuntas en el contexto de las empresas manufactureras. Sahin y Robinson (2002) revisaron la literatura a nivel operativo para estudiar el impacto de la información en tiempo real y la coordinación en la toma de decisiones con información distorsionada, variabilidad de la demanda, falta de precisión en los pronósticos, políticas de inventario y factores de inversión en un ambiente multi-eslabón y multi-periodo (Arshinder, 2008).

En forma general, según Vernadat (2002), las necesidades de integración en la cadena de suministro surgen de los problemas de integración desde diferentes perspectivas (Vernadat, 2002), (Erhan Kutanoglu, 2008):

- Integración de los mercados: Nuevas zonas comerciales y economías libres se están estableciendo en todo el mundo (Unión Europea, Mercosur, Mercado Indonesio).
- Integración entre varios sitios de fabricación y la investigación y desarrollo: Se refiere a proyectos de colaboración entre empresas que contribuyen al desarrollo de productos complejos.
- Integración entre proveedores y fabricantes: Para reducir los plazos de entrega de productos y compartir los riesgos en el lanzamiento de nuevos productos, los fabricantes y proveedores deben integrar y sincronizar sus procesos.



Tabla II. Taxonomía - Perspectivas de la coordinación de la cadena de suministro		
AUTOR (AÑO)	PERSPECTIVA	CONTEXTO
Narus y Anderson (1996)	La cooperación entre empresas independientes, pero relacionadas para compartir recursos y capacidades y satisfacer las necesidades más extraordinarias de sus clientes	Distribución de los recursos
Lambert et.al. (1999)	Un grado particular de relación entre los miembros de la cadena como un medio para compartir los riesgos y beneficios que se derivan de los resultados empresariales.	Compartir riesgos y recompensas
Ballou et.al. (2000)	La capacidad de la función logística para integrar las actividades de la cadena de suministro entre sí a través de diferentes líneas de autoridad y responsabilidad organizativa	Responsabilidad
Larsen (2000)	El trabajo colaborativo para la planificación conjunta, el desarrollo conjunto de productos, el intercambio de información mutua y sistema de información integrados, la coordinación entre diversos niveles en las empresas de la red, la cooperación a largo plazo y el reparto equitativo de los riesgos y beneficios	Visión holística de la coordinación
Lee (2000)	Coordinación de la cadena de suministro como vehículo para el rediseño de la toma de decisiones, flujo de trabajo y los recursos entre los miembros de la cadena con el objeto de mejorar el rendimiento de la cadena de suministro	Flujo de trabajo y dependencia de los recursos
Simatupang et.al. (2002)	Dada la naturaleza de las interdependencias entre las unidades que conforman la cadena de suministro, la coordinación es un requisito previo necesario para integrar sus operaciones, para lograr el objetivo común de la cadena de suministro en su conjunto	Mutualidad
Larsen et.al. (2003)	Cuando dos o más partes en la cadena de suministro planean conjuntamente una serie de actividades de promoción y elaboran proyecciones sincronizadas, sobre la base de los cuales se determinan los procesos de producción y reabastecimiento	Actividades de promoción conjuntas, pronósticos
Hill y Omar (2006)	La coordinación puede lograrse cuando los miembros de la cadena de suministro minimizan conjuntamente los costos de operación y comparten los beneficios después de planear conjuntamente la producción y definir sus políticas	Toma de decisiones conjunta y participación en los beneficios

Fuente: (Arshinder, 2008)

- Integración del diseño y la fabricación: Para reducir el tiempo de lanzamiento al mercado y reducir al mínimo los errores de diseño. Las prácticas de ingeniería concurrente se deben implementar para obtener una mayor integración entre las actividades de diseño y fabricación.
- Integración de múltiples proveedores de hardware y software: las arquitecturas de sistemas abiertos son necesarios para proporcionar la interoperabilidad de las diferentes soluciones informáticas de hardware y software de uso común en entornos industriales

Tabla III . Taxonomía - Coordinación de varias actividades e interfaces en la cadena de suministro.
Fuente: (Arshinder, 2008)

Coordinación de varias actividades e interfaces de la cadena de suministro					
Autor	Problema de coordinación	Mecanismo de coordinación	Estructura cadena suministro	Metodología usada	Medidas de rendimiento
LOGISTICA					
Stank et.al (1999)	Diferentes objetivos entre el transportista y el proveedor de transporte	Intercambio de información y tecnologías de información	Logística Proveedores y clientes	Estudio empírico	Nivel de inventarios, costos (transporte, almacenamiento, ordenar), variación ciclos de los pedidos, entregas a tiempo
Stank y Goldsby (2000)	Diferentes objetivos entre el transportista y el proveedor de transporte	Intercambio de información , alineación de metas, EDI, Contratos	Logística Proveedores y clientes	Marco conceptual	Tiempo de ciclo del canal y nivel de inventarios
Stock et.al. (2000)	Falta de integración entre logística y cadena de suministro	EDI	Logística Proveedores y clientes	Estudio empírico	Rendimiento operacional y rendimiento financiero
Huiskonen y Pirtila (2002)	Necesidad de mejorar la relación entre Logística y clientes	Intercambio de información y tecnologías de información	Logística Proveedores y manufactura	Estudio conceptual	Buenas relaciones
INVENTARIO					
Lu (1995) y Yao y Chiou (2004)	Diferentes intervalos de los pedidos	Toma de decisiones conjunta y beneficios compartidos	Un solo proveedor- múltiples compradores	Modelo analítico	Minimización de costos (ordenar, mantenimiento, compra)
Verwijmeren et.al. (1996)	Manejo independiente de los inventarios	Tecnologías de información y beneficios mutuos	Red de la cadena de suministro	Soluciones de redes	Mejora del servicio al cliente. Aumento de variedad de productos y bajos costos en la cadena de suministro
Moses y Seshadri (2000)	Necesidad de riesgo compartido. Desajuste a nivel de las existencias y el período de revisión	Toma de decisiones conjunta y descuentos por cantidad	Fabricante - Minorista	Optimización	Minimización de costos
Boyaci y Gallego (2002)	Falta de coordinación en las decisiones del tamaño del lote y el precio	Plan conjunto de precios y políticas de reposición de inventarios	Un Mayorista y múltiples minoristas	Problema analítico de optimización	Maximizar las ganancias de canal (venta al por mayor precio - inventario - gastos relacionados)



Zhao et.al. (2004)	Discrepancia en la colocación de la orden	Coordinación de pedidos, intercambio de la información	Un solo fabricante múltiples detallistas	Simulación	Minimización de costos y mejora del nivel del servicio
Wu y Ouyang (2003)	Independencia en el cálculos de los costos	Consideración de los costos conjuntos con escasez	Vendedor – comprador	Modelo analítico	Minimización de costos
Zou et.al. (2004)	Diferentes tiempos de procesamiento de pedidos de los proveedores y conflictos en los incentivos	Intercambio de información y contratos de reparto de ingresos	Múltiples proveedores – un solo fabricante o ensamblador	Modelo analítico	Maximizar utilidades
Chen y Chen (2005)	Necesidad de riesgo compartido	La consideración conjunta de los costos, el intercambio de ahorro, descuentos por cantidad	Fabricante - detallista	Matemática	Mejoramiento del pareto
Piplani y Fu (2005)	Decisiones desalineadas del inventario	Reparto de costos y contratos de nivel de servicio	Multi-eslabones	Multi-agentes tecnológicos y algoritmos genéticos	Minimización del costo de mantenimiento de inventario
Huq et.al. (2006)	Independencia en la consideración de los costos	Consideración conjunto de los costos	Multi- mayoristas Multi-detalistas	Modelo matemático y simulación	Minimización de los costos de distribución y tiempos de entrega
Barron (2007)	Diferentes tiempos de ciclo	Toma de decisiones conjunta	Cadena de suministro multieslabon	Modelo analítico	Minimización de costos (ordenar y mantenimiento)
PRONOSTICOS					
Aviv (2001)	La Toma de decisiones independiente de los pronósticos	Toma de decisiones conjunta e intercambio de información sobre la demanda	Fabricante – detallista	Modelo analítico	Minimización de costos
INVENTARIO – DISTRIBUCION					
Haq y Kannan (2006)	Mentira potencial en la reducción de costos, considerando todos los costos conjuntamente	La consideración conjunta de los costos en cada nivel	Multi-eslabón	Fuzzy AHP y algoritmos genéticos	Minimización de costos (manejo de inventario, producción, transporte)
PRODUCCION – DISTRIBUCION					
Jayaraman y Pirkul (2001)	La falta de integración en los diferentes procesos de la cadena de suministro	Minimización de los costos conjuntos de producción y distribución	Multi-eslabón	Relajación lagrangiana	Minimización de costos (compras, producción, distribución)

Pyke y Cohen (1994)	Conflicto entre el tamaño de lote grande (producción) y el tamaño de lote pequeño (distribución)	Óptimos cercanos de costo y niveles de servicio, plan conjunto	Un fabricante Un distribuidor Un detallista	Problema de optimización restringida	Costos de producción y niveles de servicio
Chandra y Fischer (1994)	Costos de manejo de inventario en múltiples locaciones de la cadena de suministro	Coordinar la programación de la producción y de rutas para vehículos	Multi-eslabón	Heurísticas	Minimización de costos (costos fijos de las instalaciones, costos de mantenimiento de inventario, costos de distribución)
Ganes han (1999)	Falta de integración en diferentes procesos	Minimización de los costos conjuntos de producción y distribución	Multi-eslabón	Programación Matemática y simulación	Minimización de costos (compras, producción, distribución)
Jang et.al. (2002)	Necesidad de coordinar producción y distribución	Minimización costo conjunto	Multi-eslabón	Heurísticas lagrangianas y algoritmos genéticos	Minimización de costos (producción, distribución)
PRODUCCION – INVENTARIO					
Yang y Wee (2002)	Conflicto en la búsqueda de número de entregas de un pedido por el vendedor y el comprador	La toma de decisiones conjunta y descuentos por cantidad	Un proveedor – múltiples compradores	Modelo matemático	Minimización de costos (mantenimiento y de ordenar)
Hill y Omar (2006)	Incremento de los costos de almacenamiento de inventario en los sucesivos eslabones de la cadena de suministro	Para diferentes actores de la cadena de suministro encontrar la cantidad optima de pedido y compartir beneficios	Un proveedor Un comprador	Modelo matemático	Minimización de costos (producción, envío y almacenamiento)
Hwang et al. (2005)	Gestión de la complejidad	Sincronización de los ciclos de producción y control de riesgos	Multi-eslabon (5 niveles)	Simulación	Nivel de existencias promedio, la cartera promedio y el costo total promedio
El cliente puede ser el proveedor, fabricante, distribuidor y detallista					

Fuente: (Arshinder, 2008)

3.3. ¿Qué tipos de integración existen?

La literatura sugiere que hay dos formas interrelacionadas de tipos de integración que los fabricantes emplean regularmente, tal como se muestra en la Figura 2.

El primer tipo de integración incluye la coordinación e integración hacia adelante del flujo físico de entregas entre proveedores, fabricantes y clientes (Saunders, 1997), (Trento y Montczka, 1998). Otros han señalado la importancia de la integración de las entregas

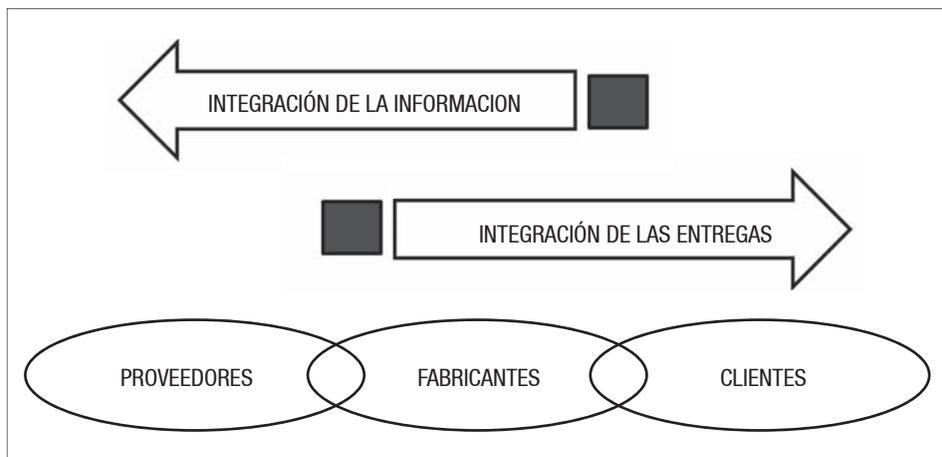


Figura 2. Integración de la cadena de suministro

en términos de implementar aplazamientos y masificación del producto en la cadena de suministro (Vernadat, 2002), (Ellinger, 1997), (Iain Dinwoodie, 2013). El otro tipo de integración prevalente incluye la coordinación hacia atrás de las tecnologías de información y el flujo de datos de los clientes a los proveedores (Vernadat, 2002), (Giuseppe Confessore, 2013).

Los flujos de información en esta dirección permitirán a las empresas responder directamente a las necesidades del cliente y puede ser coordinada a través de múltiples organizaciones o redes. Estos flujos de información fueron originalmente posibles mediante el uso del intercambio electrónico de datos (EDI). Esta tecnología ha sido sustituida por la Internet y la introducción del comercio electrónico (Marapoulos, 2005), (Tobias Schoenherra, 2012).

Las tecnologías de información permiten a múltiples organizaciones coordinar sus actividades en un esfuerzo para gestionar una verdadera cadena de suministro (Marapoulos, 2005), (Tobias Schoenherra, 2012), (Iain Dinwoodie, 2013), (Islam, 2013).

Vernadat, en su artículo “Enterprise modeling and integration”, menciona algunos tipos de integración: Integración horizontal versus vertical, Integración intra-empresa versus Integración inter-empresa, Integración de sistemas versus integración de aplicaciones versus integración de negocios (Erhan Kutanoglu, 2008), (Kar, 2013) (Leung, 2013), (Daniel Prajogo, 2012), (Marcella De Martino, 2013).

3.4. ¿Cómo se integra?

En este apartado se contemplan específicamente a los mecanismos, técnicas y herramientas que hacen posible la integración en la cadena de suministro.

Tabla IV. Taxonomía mecanismos de coordinación. Fuente: (Arshinder, 2008)

Mecanismos de Coordinación	Autores
Contratos en la cadena de Suministro	Strank y Golsby (2000), Zou et.al.(2004), Piplani y Fu (2005), Hill y Omar (2006), Yang y Wee (2002)
Tecnologías de información	Wilson (1995), Verwijmeren et.al.(1996), Lee et.al.(1999), Themistocles et.al. (2004), Caglino et.al.(2005), Fin(2005), Liu et.al.(2005)
Intercambio de información	Zouet.al.(2004), Piplani y Fu(2005), Simchi-Levi et.al.(2000), Pyke et.al.(2000), Simatupang y Sridharan(2002)
Toma de decisiones conjunta	Yang y Wee(2002), Lu(1995), Boyaci y Gallego(2002), Barron(2007), Haq y Kannan(2006), Jayaraman y Pirkul(2001), Pyke y Cohen(1993), Chandra y Fisher(1994), Ganeshan(1999), Jang et.al.(2002), Yang y Wee(2002), Hwarng et.al.(2005)

3.4.1 Mecanismos de Coordinación

El uso de mecanismos de coordinación y su correcta aplicación, permiten mejorar el rendimiento de la cadena de suministro (Arshinder, 2007). Los problemas y conflictos en la coordinación de los actores de la cadena de suministro pueden ser resueltos a través de la aplicación de mecanismos de coordinación (Jayanth Jayaram, 2010),(Arshinder, 2008). Los mecanismos de coordinación pueden ser clasificados de acuerdo a la Tabla IV, así (Arshinder, 2008):

- El uso de los cuatro mecanismos de coordinación que a continuación se explican en detalle, permiten mejorar el rendimiento de la cadena de suministro (SC) y ayudar en la evaluación de la misma (Arshinder, 2008), (Flynn, 2010), (Yao, 2010).
- Los contratos en la cadena de suministro. Los miembros de la cadena de suministro coordinan el uso de contratos para mejorar la gestión de la relación proveedor-comprador. Los contratos especifican los parámetros (por ejemplo, cantidad, precio, tiempo y calidad) en el que un comprador realiza pedidos y un proveedor los cumple. Los objetivos de los contratos de la cadena de suministro son: aumentar el beneficio total de la cadena de suministro, reducir los costos, controlar el exceso o la falta de existencias y compartir los riesgos entre los socios de la cadena de suministro (Arshinder, 2008), (Barbara B. Flynn, 2010).
- Tecnología de la información. La tecnología de la información se utiliza para mejorar la coordinación entre organizaciones (McAfee, 2002; Sanders, 2008) y, a su vez, la coordinación entre organizaciones ha demostrado que tiene un impacto positivo en seleccionar las medidas de rendimiento de la empresa, tales como servicio al cliente, tiempo de entrega y costos de producción (Vickery et.al, 2003). Las tecnologías de información (TI) ayudan a conectar el punto de producción a la perfección con el punto de entrega o compra. Permite la planeación, el requerimiento y la estimación de los plazos de entrega con base en datos en tiempo real. Los avances en TI (por ejemplo INTERNET, EDI (intercambio electrónico de datos), ERP (Enterprise resource planning), E-business (comercio electróni-



co) y muchos más), permiten que las empresas puedan intercambiar rápidamente productos, información y utilizar métodos de colaboración para optimizar las operaciones en la cadena de suministro (Vernadat, 2002), (Pagell, 2004), (Wook-Kim, 2009), (Mendoza, 2009), (Pineda, 2010), (Sanchez, 2005), (Sanchez, 2002), (Vernadat, 2002), (Jianming, 2010), (Pagell, 2004), (WookKim, 2009), (Mendoza, 2009), (Pineda, 2010), (Islam, 2013).

- El intercambio de información. Los miembros de la cadena de suministro coordinan el intercambio de información con respecto a la demanda, pedidos, inventarios, etc. El valor de intercambiar información incrementa el nivel de servicio del proveedor, reduciendo el ciclo de pedido. La política de intercambiar información produce reducciones en inventario y ahorro en costos (Yu et.al, 2001). Los datos de los puntos de venta ayudan al proveedor a anticipar mejor los futuros pedidos de los minoristas y reducen el efecto látigo (Arshinder, 2008), (Yao, 2010).
- Toma de decisiones conjunta. Las consideraciones conjuntas de reposiciones (Yao y Chou, 2004), costos de mantenimiento de inventario con demanda dinámica (Boctor et.al, 2004), planeación colaborativa (Aviv, 2001), costos de procesos diferentes (Hag y kannan, 2006; Jayraman y Pirkue, 2001; Ganeshan, 1999), frecuencia de pedidos (Yang y Wee, 2002; Burson, 2007), tamaño de lote (Pyke y Cohen, 1993; Boyaer y Gallego, 2002), desarrollo de productos (Kim y Oh, 2005), se utilizan para mejorar el rendimiento de la cadena de suministro. Una toma de decisiones coherente ayuda a resolver conflictos entre los miembros de la cadena de suministro y en el manejo de excepciones en el caso de incertidumbre (Arshinder, 2008), (Mc laren, Head, & Yufei, 2002), (Romano, 2003), (Petra Schebert, 2011), (Marcella De Martino, 2013) (Mín Huang, 2013).

Otros mecanismos de integración a coordinar que se pueden considerar son: unificación de metas, precios especiales con respecto al tamaño del lote de producción o pedido, asignación de suministro de materia prima y producto terminado por el proveedor, reducción de tiempos ciclo en el abastecimiento, aplicación de políticas de descuento, aplicación de sistemas multimodales de transporte (Rakesh Nagi, 2009), (Pamela Danese, 2013) (Pantelis Longinidis, 2013).

3.4.2 Técnicas de Integración

Marco Perona y Nicola Saccani (2004), se refieren a las técnicas y herramientas de integración particularmente en relación a la gestión de las relaciones cliente-proveedor a través de la adopción de un conjunto de prácticas que apoyan los procesos de integración. Las técnicas de integración pueden ser definidas como las decisiones sobre cómo gestionar los procesos de la integración en la cadena de suministro (Marco Perona, 2004), (Piet A. Slats, 1995), (Ramanathan, 2013). En este sentido, las técnicas de integración se pueden agrupar en tres clases (Marco Perona, 2004):

- Las técnicas para la gestión de operaciones que sirven para coordinar la logística de los socios comerciales y procesos de manufactura.

- Las técnicas para la gestión de la tecnología que sirven para coordinar e involucrar a los proveedores con los clientes en el desarrollo de procesos y actividades para nuevos productos.
- Las técnicas para planeación estratégica conjunta se dirigen a compartir la definición de objetivos de negocio y del mercado e implica directamente a la alta dirección de las empresas.

La Tabla V, proporciona una lista de técnicas de integración. A través de la adopción de técnicas de integración, las empresas pueden mejorar la eficiencia o la eficacia de los procesos de integración (Marco Perona, 2004), (Marapoulos, 2005).

Tabla V. Taxonomía Técnicas de integración. Fuente: (Marco Perona, 2004)			
Dominio	Área	Técnica	Referencias
Operaciones	Reposición	Justo a tiempo	Isaac (1985)
		Entregas frecuentes	Caputo et.al.(1996)
		Programa de reposición continua	Marien (2001)
		Inventario manejado por el vendedor (VMI)	James et.al.(1997)
	Gestión coordinada de materiales	Certificación de calidad	Manuali (1997)
		Paso libre de suministros	De Toni y Nassimbeni (1997)
	Operaciones coordinadas de planeación y control	Ordenes abiertas	Ferrozi et.al.(1993)
		Balanceo presupuestal	Jhonson (1999)
		La planificación colaborativa, previsión y reabastecimiento (CPFR)	White (2000)
	Configuración de la distribución coordinada	Reconfiguración de la red de almacenes	Stank yHaut (1990)
		Reducción de canales	Magretta (1998)
		Colocación	Bartmess y Cerny (1992)
	Gestión de la distribución coordinada	Planeación de requerimientos de distribución (DRP)	Christopher (1992)
		DRP intercompañías	Novack et.al.(1993)
		Sistemas multi-pick y multi-drop	Caputo et.al.(1996)
		Administración colaborativo del transporte (CTM)	Browning y White (2000), Cooke (2000)



Tecnología	Rediseños conjuntos	Procesos conjuntos de rediseño	Hewitt (1994)
		Diseño y rediseño de producto para la gestión de la cadena de suministro	Lee y Billington (1992)
	Desarrollo de nuevos productos	Co-diseño Ingeniería Virtual	Maggiore y Dominioni (1999)
		Innovación tecnológica conjunta	Turnbull et.al.(1992), De toni y Nassimbeni (1997), Krause (1998), Lazaric y Marengo (1997)
Planeación estratégica	Coordinación de planes estratégicos	Coordinación de negocios focalizados	Kaplan y Hurd (2002)
		Coordinación de planes de expansión de mercados	Magretta (1998)

3.4.3 Herramientas de Integración

Las herramientas de integración se pueden definir como los recursos y bienes dedicados a apoyar o permitir la adopción de una o más técnicas. Las herramientas han sido agrupadas en tres clases (Marco Perona, 2004), (Barbara B. Flynn, 2010), (Wantao Yu, 2013):

- Herramientas de información que mejoran la eficiencia y eficacia del intercambio de información en la gestión de operaciones, logística y desarrollo de nuevos productos.
- Herramientas administrativas, se utilizan para planear, medir, controlar e incentivar la ejecución de los procesos de integración.
- Herramientas organizativas, tienen el propósito de mejorar el rendimiento de los procesos de integración en los casos de interacción compleja.

La Tabla VI, muestra una lista de las herramientas de integración (Marco Perona, 2004).

3.5. ¿Para qué se integra?

Hay varias razones, entre las cuales podemos mencionar las siguientes (Damien, 2005), (Arshinder, 2008), (Laureano Jimenez, 2006), (Themistocleus, 2004), (Katrina Lintukanagas, 2009), (Pamela Danese, 2013) (Pantelis Longinidis, 2013):

- Uno de los grandes propósitos de la integración de la información es llevar a cabo la transmisión en tiempo real y procesamiento de la información necesaria para la toma de decisiones en la cadena de suministro.

Tabla VI. Taxonomía Herramientas de integración. Fuente: (Marco Perona, 2004)

Dominio	Área	Herramienta	Referencia	
Información	Contratación electrónica	Intercambio de datos electrónicos (EDI)	Banerjee y Sriram (1995), Ramamurthy et.al. (1999)	
		Mercados electrónicos (E-marketplaces)	Varda (1999)	
		Internet basados en asociaciones interactivas (I-BIP)	Ovum(1999) Economist (1999) Kehoe y Boughton (2001)	
	Integración de la información	Producción integrada y bases de datos de inventario	Hall (1997) Lee y Billington (1992)	
		Programas integrados DRP	Lee y Whang (2001)	
		Datos integrados de ingeniería	Hall (1997)	
		Administración (EDM)/producto Manejo de datos(PDM)	Marcial et.al. (1997)	
	Sistemas de monitoreo electrónico	Sistemas de monitoreo a las entregas y rastreo de los productos	Viotti (1997) Manualli (1997)	
		Sistemas de identificación automática	Lee y Whang (2001)	
	Aplicaciones de software colaborativo	Trabajo cooperativo soportado por computadora (CSCW)	Luzack y Eversheim (1999)	
		Diseño integrado asistido por computador (CAD)	Luzack y Eversheim (1999)	
		Manufactura integrada asistida por computador (CAM)	Luzack y Eversheim (1999)	
	Administrativo	Sistemas de evaluación de proveedores	Sistemas de selección de proveedores (VSS)	Masella y Rangone (2000)
			Sistemas de calificación de proveedores (VRS)	James et.al. (1997)
		Sistemas de contabilidad de costos en la cadena de suministro	Gestión de costos interorganizacional Sistema de costos kaizen	Slagmulder (2001)
Indicadores de desempeño de la cadena de suministro		Sistemas de medición del desempeño de la cadena de suministro	De Toni y Tonchia (2001) Gunasekaran et.al. (2001)	
Sistemas de incentivos en la cadena de suministro		Sistemas de incentivos en las empresas Contratos	Lee y Whang (2001) Voss y Schneidereit (2001)	



Organizativo	Roles de las interfaces	Ingenieros residentes	Caputo y Zirpoli (2001)
		Interfases con los directores	Cooper et.al. (1997)
		Alianza direccionada a las funciones gerenciales	Dyer et.al. (2001)
	Unidades organizativas en las empresas	Equipos para el desarrollo de nuevos productos	Eversheim (1996)
		Equipos de mejora de producto-proceso	Haug (1993) Handfield et.al.(2000)

- Se ha argumentado que contar con una integración logística sólida, reducirá diversos problemas, tales como el efecto látigo.
- En general, la integración logística (Integración de flujo de bienes) permite a las empresas y sus socios de la cadena de suministro actuar como una sola entidad que se traduce en un mejor desempeño a través de toda la cadena.
- De la mejora de la integración logística entre los socios de la cadena de suministro, se obtiene un número de ventajas operativas, incluido la reducción de costos (Noteboom, 1992), reducción de los plazos de entrega (Lee et.al, 2005), reducción de los riesgos (Clemens et.al, 1993), así como la mejora en las ventas, mejora en la distribución, servicio al cliente y niveles de servicio.
- La mayoría de los estudios empíricos sobre la integración de la cadena de suministro muestran una relación positiva entre la integración y el rendimiento (Van der Vaart y Van Donk, 2008). De Toni y Nassenbern (1999) encontraron que las plantas que tienen mejores rendimientos exhiben un alto nivel de interacciones logísticas, Frohlich y Westbrook (2001) encontraron que entre más amplio sea el arco de integración, mejora el rendimiento, Sheu et.al.(2006) encontró que altos niveles de colaboración resultan en eficiencias operacionales en la cadena de suministro y finalmente Li et.al (2009) encontró que la integración de la cadena de suministro está significativamente relacionada con el rendimiento de la cadena.
- El objetivo de la integración de la cadena de suministro es lograr el flujo efectivo y eficiente de productos y servicios, información, dinero y decisiones para proporcionar el máximo valor al cliente a bajo costo y alta velocidad.
- La falta de coordinación en la cadena de suministro puede resultar en un rendimiento pobre de la cadena de suministro.
- Las consecuencias de la falta de coordinación son: pronósticos inexactos, baja utilización de la capacidad, exceso de inventarios, servicio al cliente inadecuado, tiempo en el mercado, tiempo de respuesta al cumplimiento de las ordenes, calidad, atención al cliente y satisfacción del cliente.
- Existen múltiples beneficios derivados de una efectiva coordinación de la cadena de suministro (SCC). Algunos de estos beneficios incluyen: la eliminación del exceso de inventario, reducción de los plazos de entrega, aumento de las ventas, servicio al cliente, desarrollo eficiente de productos, bajos costos de fabricación,

una mayor flexibilidad para hacer frente a la incertidumbre de alta demanda, una mayor retención de clientes y mejoras en los ingresos (Fish et.al., 1994; Lee et.al., 1997; Horvath 2001). La coordinación se percibe como un requisito previo para integrar las operaciones de las entidades de la cadena de suministro para alcanzar objetivos comunes.

La Tabla VII ilustra cómo las organizaciones integran tres tipos de sistemas (Custom, Packaged y e-business solutions) que se permutan entre sí generando otros sistemas. Sobre la base de estas permutaciones, los autores definen las clasificaciones de los tipos de sistemas que se integran en la cadena de suministro (Themistocleus, 2004).

3.6 ¿Qué estrategias de integración existen?

En la última década ha habido un creciente consenso sobre la importancia estratégica de la integración de proveedores, fabricantes y clientes (Lummus et.al, 2001; Heeketal Van, 2001; Barrat, 2004). Las estrategias de integración se pueden clasificar en tres grupos: estratégicas, tecnológicas y operativas. Entre las diversas estrategias de integración, se mencionan las siguientes (Piet A. Slats, 1995), (Yuanqiong, 2012), (WookKim, 2009), (Swierczek, 2013), (Wantao Yu, 2013):

A nivel operativo las estrategias se clasifican en:

- Coordinación cliente-proveedor
- Coordinación producción-distribución
- Coordinación distribución-inventarios

A nivel estratégico, las estrategias se orientan en:

- Coordinación de la planeación estratégica
- Coordinación del enfoque de negocios

A nivel tecnológico, las estrategias se orientan en:

- Rediseño conjunto de procesos
- Desarrollo conjunto de nuevos productos
- Ingeniería virtual
- Innovación tecnológica conjunta

Entre las estrategias relacionadas con la coordinación de la gestión de materiales (a nivel operativo) se pueden mencionar entre otras:

- Adquisiciones dictaminadas
- Ordenes globales
- Dimensionamiento conjunto de la capacidad de producción
- Colaboración en la planeación de pronóstico y suministro

**Tabla VII.** Taxonomía de los diferentes tipos de sistemas que son integrados

Clasificación de tipos de sistemas	Descripción
Integración Custom-to-custom	Dentro de la organización EAI requiere la integración de aplicaciones interempresariales e intraempresariales. Al incorporar todos los sistemas requeridos, muchas aplicaciones personalizadas como aplicaciones heredadas y bases de datos están integrados en una infraestructura común, para automatizar totalmente los procesos de negocio. Como los sistemas personalizados no se han desarrollado para colaborar con otros sistemas, tienen puntos de acceso limitado a la integración. En concreto, bases de datos e interfaces de usuario son los únicos puntos posibles de integración en la mayoría de las aplicaciones personalizadas. Un caso típico de esta clasificación podría ser la incorporación de sistemas heredados que tienen que ver con la gestión de promociones (por ejemplo, existencias, cuentas de proveedores). En este caso, datos de las bases de datos e interfaces de usuario debe ser extraído y enviado de una organización (por ejemplo, minorista) a otra (por ejemplo, proveedor). Por lo tanto, las tecnologías que extraen datos de una base de datos o en una pantalla son necesarias para apoyar esta clasificación de los sistemas.
Integración Custom-to-packaged	Este es un enfoque común cuando las organizaciones adoptan EAI ya que las aplicaciones empaquetadas como los sistemas ERP en muchos casos han fracasado en lograr la integración y co-existir junto a las aplicaciones personalizadas. Un caso típico de este tipo podría ser la integración de un sistema heredado que se ocupa de la producción, y un módulo de ERP que se encarga de los pedidos de clientes o proveedores. Si bien, los sistemas ERP no fueron diseñados para incorporar otras aplicaciones autónomas, una diversidad de técnicas de enfoques y herramientas se pueden utilizar para lograr la integración entre los sistemas ERP y aplicaciones dispares. Desde una perspectiva técnica, la incorporación de sistemas ERP se puede lograr en los distintos niveles de datos, incluyendo objetos / componentes, como a nivel de interfaces.
Integración Custom-to-e-business	Muchas soluciones e-business requieren una estrecha colaboración con las aplicaciones antiguas para apoyar los procesos y tareas de los e-business habilitados. Como resultado, las aplicaciones personalizadas (por ejemplo, stocks) se incorporan con los sistemas de e-business para integrar y automatizar los procesos inter-organizacionales de negocios. Del mismo modo, en muchos casos, la funcionalidad de una solución e-business se utiliza para apoyar los sistemas personalizados. Por ejemplo, una e-store actualiza un sistema personalizado que tiene que ver con la disponibilidad de stock. La información proporcionada por la solución de e-business es fundamental no sólo para la funcionalidad de la aplicación de valores, sino también para la cadena de suministro, ya que soporta la automatización y la integración de procesos de negocio específicos. Las tecnologías que apoyan la incorporación y el intercambio de datos, objetos e interfaces son necesarias para apoyar esta clasificación de los sistemas.
Integración Packaged-to-packaged	En este caso, diferentes sistemas empaquetados, tal como las diferentes versiones de un sistema ERP o diferentes módulos de ERP que existen en una organización se unifican en una infraestructura común integrada. APIs son proporcionados por los sistemas ERP para permitir que otras aplicaciones puedan acceder a la funcionalidad ERP o de datos. Datos, mensajes y objetos se pueden introducir o transferir a un sistema ERP a través de APIs. Por lo tanto, la incorporación de un sistema empaquetado a otro sistema empaquetado requiere APIs así como también las tecnologías que soportan la extracción y transmisión de datos, mensajes y objetos.

Integración Packaged-to-e-business	Las organizaciones aprovechan la tecnología EAI y el comercio electrónico cuando integran sus soluciones e-business con paquetes de aplicaciones como sistemas ERP que pueden ser usados como sistema de back-office para apoyar la funcionalidad de e-business (aplicaciones front-end). En este caso, los procesos que tienen que ver con e-sale, e-procurement y e-supply chain se integran con los sistemas empaquetados. Aplicaciones de negocio electrónico a menudo se basan en tecnologías de objetos distribuidos (DOT) (por ejemplo, Enterprise Java Beans, CORBA, DCOM / COM) y/o lenguajes estándar orientados a Internet (por ejemplo, XML, HTML). Por lo tanto, las tecnologías que integran los datos, los objetos, las interfaces y los mensajes son importantes para la integración de los sistemas empaquetados y la integración de e-business.
Integración Ebusiness-to-business	En este enfoque, una aplicación de e-business está integrado y es compatible con la funcionalidad de otra solución de e-business. Por ejemplo, un punto de venta electrónico es incorporado con e-supply chain para compartir datos que son importantes para la aplicación de esta última (por ejemplo, pedidos de clientes, datos de cliente, etc.) La integración de aplicaciones e-business puede ser facilitado por las tecnologías basadas en mensajes (por ejemplo, XML), distribuidos tecnologías de objetos (por ejemplo, CORBA) y tecnologías de bases de datos orientados (Java Database Connectivity, JDBC).
Integración Custom-to-packaged-to-e-business	Este enfoque se centra en el desarrollo de una infraestructura integrada que integra los procesos y aplicaciones en los departamentos, empresas o a nivel de toda la empresa. Los tipos de sistemas que se incorporan requieren tecnologías de integración que soportan todos los niveles de integración. Por lo tanto, las tecnologías que faciliten datos, objetos, interfaces y mensajes son requeridas.

Fuente: (Themistocleus, 2004).

Las estrategias más comunes para el diseño de la red de distribución (a nivel estratégico) son:

- La reconfiguración de la red de almacenes
- Acortamiento del canal de comercialización
- Localización.

Las estrategias de gestión de distribución (a nivel operativo) incluyen:

- Coordinación de la planeación de los requerimientos de distribución entre compañías.
- Sistemas multiloteo (multi-pick)
- Multientregas (Multi-drop)
- Esquemas de gestión de transporte colaborativo.

Dentro de las estrategias de coordinación para la gestión de inventarios (a nivel operativo), podemos mencionar:

- Estrategia de desarrollo conjunto de ordenes (DCO)
- Estrategias justo a tiempo
- Respuesta rápida (QR: Quick Response)
- Estrategia de reaprovisionamiento eficiente (ER)
- Reaprovisionamiento continuo (CR)



- Planeación, Pronostico y reabastecimiento (CPFR)
- Inventario administrado por el proveedor (VMI)
- Estrategia gestión de disponibilidad por el proveedor (SMA)
- Estrategia épocas comunes de resurtido (CRE)

3.7 ¿Cómo se mide la integración?

Muchos autores parecen estar de acuerdo con que un elevado nivel de integración tiene un impacto positivo sobre los resultados de las empresas pertenecientes a la cadena de suministro. Mientras algunos autores han aportado evidencia empírica de esta afirmación (Frohlich y Westbrook, 2001; Droge et.al., 2004; Gimenez y Ventura, 2005), otros tienen dudas y sugieren que esta relación es más retórica que real (Van der Vaart y Van Donk, 2008; Fabbe-Costes y Jahre, 2007) ya que existe poco consenso en cómo capturar la esencia de la integración y que los tipos de resultados sobre los que se mide su influencia varían de unos autores a otros (Romano, 2001), (Frohlich, 2001), (Paul Cousins, 2006), (WookKim, 2009), (Qiu Hong Zhao S. C., 2010), (Yan Cui, 2013), (Ahmed Musa, 2013).

En la literatura, para medir la integración se proponen diversos indicadores: la intensidad de la comunicación, la coordinación de actividades, la creación de equipos, la ausencia de fronteras rígidas entre las actividades logísticas de las empresas y las de los proveedores o clientes, la elaboración de decisiones conjuntas o el intercambio de información (Thomas y Griffin, 1996; O'Leary-Kelly y Flores, 2002). Pero, ¿por qué las empresas realizan estos cambios en su estructura y en su relación con otros agentes de la cadena de suministro? (Frohlich, 2001), (Paul Cousins, 2006), (WookKim, 2009), (Nagurney, 2009), (Devendra Choudhary, 2013).

La respuesta es que existe un consenso de que a mayor grado (intensidad) de integración, la empresa va a obtener una ventaja competitiva y con ella, unos mejores resultados. Los resultados logísticos más estudiados han sido la eficiencia en costes, la fiabilidad de entrega y la flexibilidad. En primer lugar, la integración logística supondrá para la empresa un ahorro en costos por varios motivos (Uk Mike, 1996), (Frohlich, 2001), (Marc, 2002), (Paul Cousins, 2006), (Taco Van Der Vaart, 2008), (WookKim, 2009), (Cornelia Droge, 2012), (Danese, 2013), (Islam, 2013):

- La reducción de inventario, ya que la relación con proveedores hace posible acuerdos de entrega de materias primas que favorecerán la reducción del stock,
- La disminución de los costes de transacción y de la incertidumbre de previsión de la demanda y del aprovisionamiento y
- Una mejor coordinación y cooperación entre las áreas funcionales y entre las empresas participantes.

En segundo lugar mejora la fiabilidad de entrega, es decir, la capacidad que tiene una organización de proveer a tiempo el tipo y la cantidad de producto requerido por

los clientes en cada momento (Frolich, 2001), (Paul Cousins, 2006), (WookKim, 2009), (Giuseppe Confessore, 2013).

En tercer lugar, aumenta la flexibilidad, que es la capacidad del sistema logístico para comunicarse con los clientes y ajustar los esquemas de producción y distribución según las necesidades particulares de los mismos (Frolich, 2001), (Paul Cousins, 2006), (WookKim, 2009), (Chee Yew Wonga, 2011).

Frohlich y Westbrook (2001) dieron respuesta a esas cuestiones (intensidad y dirección de la integración) mediante un instrumento denominado *arcos de integración* donde se muestra cómo las empresas toman decisiones estratégicas respecto a la amplitud deseada de su integración y la dirección de la misma: aguas arriba o aguas abajo. Los distintos arcos de integración se construyen siguiendo las instrucciones que aparecen en la Figura 3 (Frolich, 2001), (GunaseKaran, 2004), (Muñoz, 2006), (Yuliang Yaoa, 2007), (Xiande Zhao, 2008), (Gang Li, 2009), (Bernd Sholz-Reiter, 2010), (Danese, 2013).

Como se observa en la Figura 3, algunos fabricantes deciden participar relativamente poco con proveedores y clientes y así tener un arco relativamente estrecho de integración. Otros fabricantes, al contrario, aplican una estrategia con un amplio arco de integración, es decir, integran ampliamente sus organizaciones con los proveedores y los clientes. Es evidente que a un mayor nivel de integración entre proveedores y clientes de la cadena de suministro, mayores serán los beneficios (Stevens, 1989; Lee et.al, 1997; Metters, 1997; Naransinham y Jayaran, 1998; Anderson y Katz, 1998; Hines et.al, 1998; Jhonson, 1999). Tan et.al (1998), señalaron que cuando las empresas se integran y actúan como una entidad, el rendimiento se mejora en toda la cadena. Handfield y Nichols (1999), argumentaron que ahora los fabricantes no solo deben gestionar sus propias organizaciones sino también estar involucrados en la gestión de la red, tanto en los flujos de información como en los flujos de bienes físicos a través de la cadena de suministro (Frolich, 2001).

El nivel de integración se mide a través del análisis factorial sobre actividades de colaboración entre la empresa y sus clientes y proveedores. La puntuación factorial obtenida por cada empresa (método de regresión) sirve para clasificarla en el cuartil superior (superior al 75%), inferior (inferior al 25%) o intermedio. Así se definen cinco grupos, excluyentes entre sí: empresas con bajos niveles de integración con proveedores y clientes (arco interno), aquellas con niveles intermedios (arco periférico) y aquellas con niveles elevados (arco externo). Cuando hay asimetría y se está integrado con un miembro del canal y no con el otro, da lugar al arco hacia el proveedor y el arco hacia el cliente. Las conclusiones a las que llegaron Frohlich y Westbrook (2001) fueron que a mayor amplitud del arco, mejores resultados y que si el arco era asimétrico los resultados no eran mejores que si no había integración (Marapoulos, 2005), (Sabine Fliess, 2006), (Qiu Hong Zhao S. C., 2010), (Salih Zeki, 2011), (Tobias Schoenherr, 2012), (Yuanqiong, 2012), (Farahani, 2013).



¿HASTA QUÉ PUNTO INTEGRA SU ORGANIZACIÓN SUS ACTIVIDADES CON PROVEEDORES Y CLIENTES CLAVE?	TIPOS DE ARCOS DE INTEGRACIÓN
<p>Arco Interno</p> <p>Están clasificados como ARCO INTERNO aquellas empresas cuyo nivel de integración fue:</p> <ul style="list-style-type: none"> a.- Poco integrado con proveedores (en el cuartil inferior de las respuestas para proveedores) b.- Poco integrado con clientes (en el cuartil inferior para clientes) 	
<p>Arco Periférico</p> <p>Están clasificados como ARCO PERIFÉRICO aquellas empresas cuyo nivel de integración fue:</p> <ul style="list-style-type: none"> a.- Medianamente integrado con proveedores (por encima del cuartil inferior y por debajo del superior) b.- Medianamente integrado con clientes (por encima del cuartil inferior y por debajo del superior) 	
<p>Arco Externo</p> <p>Están clasificados como ARCO EXTERNO aquellas empresas cuyo nivel de integración fue:</p> <ul style="list-style-type: none"> a.- Altamente integrado con proveedores (en el cuartil superior) b.- Altamente integrado con clientes (en el cuartil superior) 	
<p>Arco Hacia Proveedores</p> <p>Están clasificados como ARCO HACIA PROVEEDORES aquellas empresas cuyo nivel de integración fue:</p> <ul style="list-style-type: none"> a.- Altamente integrado con proveedores (por encima del cuartil superior) b.- Medianamente o poco integrado con clientes (por debajo del cuartil superior para clientes) 	
<p>Arco Hacia Clientes</p> <p>Están clasificados como ARCO HACIA CLIENTES aquellas empresas cuyo nivel de integración fue:</p> <ul style="list-style-type: none"> a.- Medianamente o poco integrado con proveedores (por debajo del cuartil superior) b.- Altamente integrado con clientes (por encima del cuartil superior) 	

Figura 3. Arcos de Integración. Fuente: (Frolich, 2001)

La relación de indicadores de desempeño adecuados es un reto, debido a la complejidad inherente y la interdependencia de las cadenas de suministro. Mientras Chen y Paulraj (2004) sostiene que el desempeño financiero debe ser el criterio principal de valoración del desempeño de la cadena de suministro, otros han descrito las limitaciones de confiar únicamente en las medidas financieras de rendimiento (Dixon et.al, 1990, Eccles y Burn, 1992, Hall, 1993. Jhonson y Kaplan, 1987; Skinner, 1971). Van Hock (1998) y Beamon (1999) sugirieron que la medición del desempeño de la cadena de suministro

debe incluir indicadores operacionales, tales como servicio al cliente y la capacidad de responder a un entorno cambiante. Nelly et.al (1995) enumeran costo, tiempo, calidad, entrega y flexibilidad como medidas importantes del rendimiento operacional (Themistocleus, 2004), (Indranil Bose, 2008), (Christos Maravelias, 2009), (Burcu Keskin, 2010), (Jiang Yanhua, 2011), (C.F. Cheung, 2012), (Devendra Choudhary, 2013).

De otro lado, los principales indicadores que caracterizan la competitividad del sistema logístico y los niveles que en ellos muestran las empresas líderes en países desarrollados como Japón, Francia, Estados Unidos o Alemania son (Acero E, 2012), (Edrisi Muñoz, 2013), (Farahani, 2013):

- Inventario promedio/Ventas = 20,0 %
- Costo logístico/Ventas = 5,1 %
- Oportunidad en los aprovisionamientos = 90,0 %
- Oportunidad en los suministros a los clientes contra pedidos = 90,0 %
- Oportunidad en los suministros a los clientes contra almacén = 88,0 %
- Suministros perfectos de los proveedores = 89,5 %
- Suministros de pedidos perfectos a los clientes = 95,0 %
- Utilización de las capacidades de producción = 80,0 %
- Utilización de las capacidades de almacenaje = 90,0 %
- Cobertura del inventario de productos terminados = 21,0 días
- Cobertura del inventario de materia prima y materiales = 43,5 días

Las empresas tienen definido un sistema de indicadores para caracterizar y evaluar la gestión logística en general y en cada una de las actividades y unidades que conforman el sistema logístico de la organización, por otra parte, realizan sistemáticamente Benchmarking con las empresas que exhiben resultados de avanzada en las distintas actividades logísticas. Además, tiene adecuadamente organizado todos los registros de dichos indicadores. Dichos registros sirven de fundamento para una activa gestión de mejora del servicio al cliente sobre la base de lograr una adecuada diferenciación de los clientes. En última instancia, la efectividad de la gestión del sistema logístico tiene como fin elevar metódicamente la competitividad de la empresa. En tal caso la empresa exhibe el impacto del sistema logístico en un crecimiento sostenido de la competitividad de la compañía, la cual se refleja en la satisfacción sistemática de los siguientes indicadores (Diana Ballesteros, 2004), (Giuseppe Confessore, 2013), (Islam, 2013):

- Ritmo de crecimiento de los ingresos anuales
- Un ritmo de crecimiento de la ganancia superior al ritmo de crecimiento de los ingresos
- Crecimiento del mercado objetivo, incluyendo el aumento de las exportaciones
- Aumento de la cuota de mercado
- Crecimiento de la productividad y del salario medio.



4. Conclusiones

En términos generales, la integración de toma de decisiones en la red logística requiere tiempo, recursos, inversión, infraestructura, apropiación de tecnologías de información para ser abordada en forma efectiva. En cuanto a las tecnologías de información, es importante resaltar que requieren ajustes e integración con otras plataformas, específicamente el software desarrollado como ERP. Esto hace que la implementación ERP sea compleja y requiera el desarrollo de sistemas de información para mejorar los procesos de integración.

En la revisión realizada se evidenció que aún cuando se han desarrollado e implementado procesos de integración con resultados positivos, no obstante todavía falta por mejorar la integración de la toma de decisiones en la red logística y para ello se siguen desarrollado modelos matemáticos, software, sistemas de información, estrategias de integración para mejorar aun mas los procesos de integración . Algunas conclusiones adicionales como resultado de la revisión realizada, son:

- El concepto de integración de la cadena de suministro, pretende englobar todos los actores (proveedores, productores, distribuidores y clientes) de la cadena de suministro trabajando en forma conjunta, con objetivos y decisiones comunes, considerando la existencia de un único flujo de productos, y analizando y mejorando los procesos entre cada uno de los actores con el fin de aportar un beneficio global a toda la cadena de suministro.
- La falta de integración en la cadena de suministro genera pronósticos inexactos, baja o excesiva utilización de la capacidad de producción, desconocimiento de la capacidad requerida, exceso o falta de inventarios, efecto látigo, servicio al cliente inadecuado, tiempo de respuesta tardíos al cumplimiento de las ordenes, costos de transporte elevados, costos de fabricación elevados, costos de inventario elevados, errores de envío, errores en la recepción, errores de facturación/pagos, rendimiento pobre de la cadena.
- El uso de tecnologías de información (E-commerce, E-Business, B2B, B2C, EAI, ERP) permiten mejorar la integración de la cadena de suministro.
- El uso de Outsourcing de servicios (Empresas consultoras), tales como 3PLs, 4PLs, 5PLs y 7PLs, facilitan la integración en la cadena de suministro.
- La integración de la cadena de suministro y las tecnologías de la información es una parte esencial del éxito de cualquier compañía que quiera sobrevivir en el mercado actual. Esta combinación permite tomar decisiones con mayor facilidad y rapidez.

Referencias Bibliográficas

- [1] Acero E, M. (2012). Administración de la cadena de suministro. *Series de gerencia Empresarial*.
- [2] Ahmed Musa, A. G. (2013). Embedded devices for supply chain applications: Towards hardware integration of disparate technologies. *Expert Systems with Applications*.
- [3] Arshinder, A. K. (2008). Supply chain coordination: Perspectives, empirical studies and research directions. *Int.J. Production Economics*, 316-335.
- [4] Barbara B. Flynn, B. H. (2010). The impact of supply chain integration on performance. *Journal of Operations Management*, 58-71.
- [5] Benjamin, J. (1989). An Analysis of inventory and transportation cost in a constrained network. *Transportation Science. Volumen 23*, 177-183.
- [6] Bernd Sholz-Reiter, E. M. (2010). Integrating manufacturing and logistic system along global supply chains. *Journal of Manufacturing Science and Tecnology*, 216-223.
- [7] Blumenfeld, B. D. (1985). Analizing trade-offs between transportation, inventory and production costs on freight networks. *Transportation Research*, 361-380.
- [8] Blumenfeld, B. D. (1991). Synchronizing production and transportation schedules. *Transportation Research Part B Volumen 25*, 23-37.
- [9] Bookbinder, Q. I. (1999). An integrated inventory-transportation system with modified periodic policy for multiple products. *European Journal Operation Research, Volumen 115*, 245-269.
- [10] Brent D. Williams, J. R. (2013). Leveraging supply chain visibility for responsiveness: The moderating role of internal integration. *Journal of Operations Management*.
- [11] Burcu Keskin, H. U. (2010). Integration of strategic and tactical decisions for vendor selection under capacity constraints. *Computers & Operations Research*, 2182-2191.
- [12] C.F. Cheung, S. K. (2012). A Knowledge-based customization system for supply chain integration. *Expert Systems with Applications*, 3906-3924.
- [13] Cesar, P. (2010). Aplicaciones computacionales en la cadena logistica integrada.
- [14] Chee Yew Wonga, S. B. (2011). The contingency effects of environment uncertainty on the relationship between supply chain integration and operational performance. *Journal of Operations Management*, 604-615.
- [15] Chien. (1993). Determining profit-maximizing production/shipping policies in a one-to-one direct shipping, stochastic demand environment. *European Journal Operations Research, Volumen 64*, 83-102.
- [16] Christina W.Y. Wong, C. Y. (2013). The combined effects of internal and external supply chain integration on product innovation. *Int. J. Production Economics*.
- [17] Christos Maravelias, C. S. (2009). Integration of production planning and scheduling. *Computers and Chemical Engineering*, 1919-1930.
- [18] Claudia Colicchia, G. M. (2013). Building environmental sustainability: empirical evidence from Logistics Service Providers. *Journal of Cleaner Production*.
- [19] Cohen, L. (1988). Strategic analysis of integrated production-distribution systems: Models and methods. *Operation Research, Volumen 36*, 217-227.
- [20] Cornelia Droge, S. V. (2012). Does supply chain integration mediate the relationships between product/process strategy and service performance. *Int.J. Production Economics*, 250-262.



- [21] Damien. (5 de 15 de 2005). Strategy development processes as determinants of B2b e-commerce performance: A comparative Power. *Proquest Education Journals*.
- [22] Danese, P. (2013). Supplier integration and company performance: A conFIGurational view. *Omega*, 93.
- [23] Daniel Prajogo, J. O. (2012). Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration. *Int.J. Production Economics*, 514-522.
- [24] Devendra Choudhary, R. S. (2013). Joint decision of procurement lot-size, supplier selection, and carrier selection.
- [25] Diana Ballesteros, P. B. (2004). La logística competitiva y la administración de la cadena de suministros. *Scientia et Technica Año X No 24*.
- [26] Edrisi Muñoz, E. C.-G. (2013). Integration of enterprise levels based on an ontological framework. *Chemical Engineering Research and Design*.
- [27] Ellinger, A. (1997). The relationship between integrated logisitcs and customer service.
- [28] Erhan Kutanoglu, D. L. (2008). Integrated inventory and transportation mode selection: A service parts logistics system. *Transportation Research*, 665-683.
- [29] Farahani, R. Z. (2013). Competitive supply chain network design: An overview of classification models, solution techniques and applications.
- [30] Ferdinand Jaspers, J. V. (2006). The organizational form of vertical relationship: Dimensions of integration. *Industrial Marketing Management*, 819-828.
- [31] Frank Wiengarten, M. P. (2013). Do a country's logistical capabilities moderate the external integration performance relationship? *Journal of Operations Management*.
- [32] Frolich, W. (2001). Arcs of integration: an international study of supply chain strategies. *Journal of Operations Management*, 185-200.
- [33] Fumero, V. (1997). Integrating distribution, machine assignment and lot sizing via Lagrange Relaxation. *International Journal Production Economic*, volumen 49, 45-54.
- [34] Gang Li, H. Y. (2009). The impact of IT implementation on supply chain integration and performance. *Int.J. Production Economics*, 125-138.
- [35] Giuseppe Confessore, G. G. (2013). A production and logistics network model with multi-modal and sustainability considerations.
- [36] GunaseKaran. (2004). Information systems in supply chain integration and management. *European Journal of Operation Research*, 269-265.
- [37] Hall. (1996). On the integration of production and Distribution: Economic order and production quantity implications. *Transportation Research, Part B volumen 30*, 387-403.
- [38] Iain Dinwoodie, D. M. (2013). Development of a Combined Operational and Strategic Decision Support Model for Offshore Wind. *Energy Procedia*.
- [39] Indranil Bose, R. P. (2008). ERP and SCM system integration. *Information and Management*, 233-241.
- [40] Islam, D. M. (2013). Logistics and supply chain management. *Research in Transportation Economics*.
- [41] J.F., W. (1981). Heuristics techniques for simultaneous scheduling of production and distribution in multi-echelon structures. *Management Sciences*, 336-352.
- [42] Jayanth Jayaram, k.-C. T. (2010). Supply chain integration with third-party logistics providers. *Int.J. Production Economics*, 262-271.

- [43] Jiang Yanhua, L. X. (2011). A study on supply chain information integration of commodity circulation based on grid. *Engineering*, 553-557.
- [44] Jianming, Y. (2010). Decision optimization analysis on supply chain resource integration in fourth party logistics. *Journal of Manufacturing Systems*, 121-129.
- [45] Kar, A. K. (2013). Revisiting the supplier selection problem: An integrated approach for group decision support.
- [46] Katrina Lintukangas, S. (2009). Some issues of supply chain management integration. *Journal of Purchasing & Supply Management*, 240-248.
- [47] Laureano Jimenez, R. M. (2006). Integration of supply chain management and logistics. *Elsevier B.V.*, 25-35.
- [48] Leung, L. C. (2013). Managing third-party logistics under uncertainty: A decision scheme and managerial implications. *Int. J. Production Economics*.
- [49] Maropoulos, K. a. (2005). A theoretical Framework for the Integration of resource aware planning with logistics for the dynamic validation of aggregate plan within a production network. *University of Bath*, 33-36.
- [50] Marcella De Martino, L. E. (2013). Logistics innovation in seaports: An inter-organizational perspective. *Research in Transportation Business & Management*.
- [51] Marco Perona, N. S. (2004). Integration Techniques in customer-supplier relationship. *Int.J. Production Economics*, 189-205.
- [52] Mc laren, T., Head, M., & Yufei, Y. (4 de 12 de 2002). Supply chain collaboration alternatives: Understanding the expected costs and benefits. *Proquest Education Journal*, 348.
- [53] Mendoza, A. (2009). The Total Cost of Logistic in Supplier Selection Decisions. *Proceeding of the 2009 Industrial Engineering Research Conference*, 47-56.
- [54] Mendoza, A. (2009). The total cost of logistic in supplier selection decisions.
- [55] Min Huang, Y. C. (2013). Fourth party logistics routing problem with fuzzy duration time.
- [56] Mithun Sharma, I. M. (2008). 2008. *Applied mathematics and computation*, 256-265.
- [57] Muñoz, L. J. (2006). Integration of supply chain management and logistics: development of an electronic data interchanges for SAP servers. *Elsevier B.V.*, 23-27.
- [58] Nagurney, A. (2009). A system-optimization perspective for supply chain network integration: The horizontal merger case. *Transportation Research*, 1-15.
- [59] Neal Juster, A. d. (1997). Enterprise modeling and integration: a taxonomy of seven key aspects. *Computer and Industry*, 339-359.
- [60] Pagell, M. (2004). Understanding the factors that enable and inhibit the integration of operations, purchasing and logistics. *Journal of Operations Management*, 459-487.
- [61] Pamela Danese, P. R. (2013). The impact of supply chain integration on responsiveness: The moderating effect of using an international supplier network. *Transportation Research Part E*.
- [62] Pantelis Longinidis, M. G. (2013). Integration of sale and leaseback in the optimal design of supply chain networks.
- [63] Paul Cousins, B. M. (2006). The implications of socialization and integration in supply chain management. *Journal of Operations Management*, 604-620.
- [64] Petra Schebert, C. L. (2011). B2B integration in global supply chains: An identification of technical integration scenarios. *Journal of Strategic Information Systems*, 250-267.



- [65] Piet A. Slats, B. B. (1995). Modeling, Logistic chain. *European Journal of Operational Research*, 1-20.
- [66] Pineda, C. (2010). Aplicaciones computacionales en la cadena logística integrada. *Tesis de grado Especialización Ingeniería de Producción*.
- [67] Qiu Hong Zhao, S. C. (2010). Integration of inventory and transportation decisions in a logistics system. *Transportation Research*, 913-925.
- [68] Qiu Hong Zhao, S. C. (2010). Integration of inventory and transportation decisions in a logistics system. *Transportation Research*, 913-925.
- [69] Rakesh Nagi, F. P. (2009). Multi-echelon supply chain network design. *Proceeding of the 2009 Industrial Engineering Research Conference*, 30-35.
- [70] Ramanathan, U. (2013). Performance of supply collaboration - A simulation study.
- [71] Romano, P. (2001). Co-ordination and integration mechanisms to manage logistics processes across supply networks. *Pergamon*, 23.
- [72] Romano, P. (2003). Co-ordination and integration mechanisms to manage logistics processes across supply networks. *Journal of Purchasing & Supply Management*, 119-134.
- [73] Sabine Fliess, U. B. (2006). Supplier integration - controlling of co-development processes. *Industrial marketing Management*, 28-44.
- [74] Salih Zeki, H. K. (2011). The effect of supply chain integration on information sharing. *Procedia Social and Behavioral Sciences*, 1630-1649.
- [75] Sanchez, J. E. (2002). *Marco Conceptual de la cadena de suministro*. México: sanfandila, nro 215.
- [76] Sánchez, J. E. (2005). *Estado del Arte de los modelos matemáticos para la coordinación de inventarios en la cadena de suministro*. México: Sanándola No 281.
- [77] Sánchez, J. E. (2006). *Coordinación de inventarios en una cadena de suministro del sector automotriz*. Mexico: Sanfandila, nro 293.
- [78] Sharp, S. G. (1989). A decomposition algorithm for solving the multi-facility production-transportation problem with nonlinear production costs. *Econometric, Volumen 38*, 450-506.
- [79] Speranza, U. (1994). Minimizing transportation and inventory costs for several products on a single link. *Operation Research, Volumen 42*, 879-893.
- [80] Swierczek, A. (2013). The impact of supply chain integration on the “snowball effect” in the transmission of disruptions: An empirical evaluation of the model. *Int. J. Production Economics*.
- [81] Systems, M. a. (2002). Marc Goetschalckx, Carlos Vidal. *European Journal of Operational Research*, 1-18.
- [82] Taco Van Der Vaart, D. P. (2008). A critical review of survey-based research in supply chain integration. *Int.J. production Economics*, 42-55.
- [83] Taskin Gumus, F. G. (2007). Multi-echelon inventory management in supply chain with uncertain demand and lead times: literature review from an operational research perspective. *European Journal of Operational Research*, 225-239.
- [84] Themistocleus, Z. I. (2004). Evaluating the integration of supply chain information systems: A case study. *European Journal of Operational Research*, 393-405.
- [85] Thomas, G. (1996). Coordinated supply chain management. *European Journal Operation Research, Volumen 94*, 1-15.
- [86] Tobias Schoenherr, M. S. (2012). Revisiting the arcs of integration: Cross-validations and extensions. *Journal of Operations Management*, 99-115.

- [87] Uk Mike. (1996). Integrated Logistics Call in the Revolutionaries.
- [88] Vernadat, F. (2002). Enterprise modeling and integration. *Annual reviews in control* 26, 15-25.
- [89] Vijayasathy, L. R. (2010). Supply integration: An investigation of its multi-dimensionality and relational antecedents. *Int.J. Production Economics*, 489-950.
- [90] Wang Sena, S. P. (2004). Supply chain positioning strategy integration, evaluation, simulation and optimization. *Computers & Industrial Engineering*, 781-792.
- [91] Wantao Yu, M. A. (2013). The effects of supply chain integration on customer satisfaction and financial performance: An organizational learning perspective. *Int. J. Production Economics*.
- [92] Wei Zhang, D. X. (2013). Integrating the logistics network design with order quantity determination under uncertain customer demands. *Expert Systems with Applications*.
- [93] WookKim, S. (2009). An investigation on the direct and indirect effect of supply chain integration on firm performance. *Int.J. Production Economics*, 328-346.
- [94] Xiande Zhao, B. H. (2008). The impact of power and relationship commitment on the integration between manufacturers and customers in a supply chain. *Journal of Operation Management*, 368-388.
- [95] Xiande Zhaoa, B. H. (2011). The impact of internal integration and relationship commitment on external integration. *Journal of Operations Management*, 17-32.
- [96] Xu, J. (2013). Integrated inventory management and supplier base reduction in a supply chain with multiple uncertainties. *European Journal of Operational Research*.
- [97] Yan Cui, M. H. (2013). Fourth party logistics routing problem model with fuzzy duration time and cost discount. *Knowledge-Based Systems*.
- [98] Yao, J. (2010). Decision optimization analysis on supply chain resource integration in fourth party logistics. *Journal of Manufacturing Systems*, 121-129.
- [99] Yuanqiong. (2012). Supply chain integration service oriented transformation evidence from Chinese equipment manufacturers. *Int.J. Production Economics*, 791-799.
- [100] Yuliang Yao, P. E. (2007). Supply chain integration in vendor-managed inventory. *Decision Support Systems*, 663-667.

Cesar Augusto Pineda Pérez

Magister en Ingeniería Industrial, Especialista en Ingeniería de Producción e Ingeniero Industrial de la Universidad Distrital Francisco José de Caldas de Bogotá. . Experiencia en diversas empresas del sector industrial. Se ha desempeñado como docente en diversas universidades públicas y privadas en el área de investigación de operaciones, producción y logística.

Cesar Amílcar López Bello

Magister en Ingeniería Industrial, Universidad de los Andes. Especialista en Ingeniería de Producción, Universidad Distrital. Ingeniero Industrial, Universidad Distrital. Profesor investigador grupo sistemas logísticos de la Universidad de la Sabana. Profesor Asociado, Facultad de Ingeniería Universidad Distrital. Investigador grupo MMAI de la universidad distrital.

Instrucciones para los Autores

1. INTRODUCCIÓN

La revista INGENIERÍA es una publicación periódica-científica de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas. Tiene como propósito principal divulgar los resultados de investigación de la comunidad académica de la Facultad del país y del exterior. Por tal motivo, los artículos que se publican en la revista deben tener relación con el trabajo en ingeniería, en general, y con las derivaciones de la ingeniería que se desarrollan en la Universidad, en particular. Su frecuencia es semestral (dos ediciones por año). Todo artículo es evaluado de manera doble-ciega por pares revisores antes de ser calificado como candidato para publicar. El Comité Editorial, tomando en cuenta los conceptos emitidos por los evaluadores y el cumplimiento de las políticas editoriales de la Revista, decide la publicación del artículo en una edición futura. Todos los artículos presentados a la revista INGENIERÍA deben ser inéditos. No se publicarán artículos que hayan sido anteriormente publicados o que hubieren sido sometidos por el autor o los autores a otros medios de difusión nacional o internacional para su publicación. Los artículos se publican conforme a la tipología definida del Departamento Administrativo de Ciencia, Tecnología COLCIENCIAS, para revistas indexadas (*Artículo de investigación científica y tecnológica, Artículo de reflexión, Artículo de revisión, Artículo corto, Reporte de caso, Revisión de tema*). Se le dará prelación a los artículos de investigación científica y tecnológica, reflexión o revisión, en las diferentes ramas de la ingeniería.

2. CLASIFICACIÓN DE LOS ARTÍCULOS

La clasificación de los artículos dentro de la revista INGENIERÍA se hace de acuerdo con la categorización sugerida por el Instituto Colombiano para el Desarrollo de la Ciencia y la Tecnología, COLCIENCIAS, para revistas científicas y tecnológicas indexadas. Los autores deben determinar a cuál de las categorías pertenece el artículo que están enviando, en la carta de presentación. (Véase la sección 3.1.)

Las categorías de los artículos están identificadas por el número de orden de la siguiente lista y sus características son:

- **Artículo de investigación científica y tecnológica.** Documento que presenta, de manera detallada, los resultados originales de

proyectos de investigación. La estructura generalmente utilizada contiene cuatro apartes importantes: introducción, metodología, resultados y conclusiones.

- **Artículo de reflexión.** Documento que presenta resultados de investigación desde una perspectiva analítica, interpretativa o crítica del autor, sobre un tema específico, recurriendo a fuentes originales.
- **Artículo de revisión.** Documento resultado de una investigación donde se analizan, sistematizan e integran resultados de investigaciones publicadas o no publicadas, sobre un campo en ciencia o tecnología, con el fin de dar cuenta de los avances y las tendencias de desarrollo. Se caracteriza por presentar una cuidadosa revisión bibliográfica de por lo menos 50 referencias.
- **Artículo corto.** Documento breve que presenta resultados originales preliminares o parciales de una investigación científica o tecnológica, que por lo general requieren de una pronta difusión.
- **Reporte de caso.** Documento que presenta los resultados de un estudio sobre una situación particular con el fin de dar a conocer las experiencias técnicas y metodológicas consideradas en un caso específico. Incluye una revisión sistemática comentada de la literatura sobre casos análogos.
- **Revisión de tema.** Documento resultado de la revisión crítica de la literatura sobre un tema en particular.

La revista INGENIERÍA dará prelación a aquellos documentos que se clasificarían en los artículos tipo 1, 2 ó 3. El interés fundamental está en la divulgación del trabajo científico realizado en ingeniería, en general, y en la Facultad de Ingeniería de la Universidad Distrital, en particular.

3. PRESENTACIÓN DEL DOCUMENTO

La plantilla para la elaboración de artículos en formato .doc puede descargarse del portal Web de la Revista en la siguiente dirección:

<http://revistas.udistrital.edu.co/ojs/index.php/reving/article/view/4979/6709>

Los artículos deben ser escritos en hojas de papel de tamaño carta estándar (21,59 cm × 27,94 cm), en una sola columna. El tipo de letra en las columnas del artículo es Times New Roman o Calibri de tamaño 10. El espacio entre líneas es de uno (1) o espaciado sencillo. Los párrafos inician en la primera línea con una sangría de 0.3 cm.

La primera página del artículo debe contener título, resumen y palabras claves en español y en inglés, además de los datos de contacto de los autores. El título no debe contener más de 20 palabras y no debe incluir abreviaturas, paréntesis, fórmulas, siglas o caracteres desconocidos. Debe ser breve, preciso y codificable de manera que pueda ser registrado en índices internacionales. El Resumen debe tener una ex-

tensión máxima de 200 palabras, sin abreviaturas, ni referencias bibliográficas, ni caracteres desconocidos. En él se señala en forma concisa los objetivos, hallazgos y conclusiones del estudio descrito en el artículo. El resumen debe ser comprensible sin necesidad de recurrir al resto del texto, figuras o tablas del artículo. Debe ir acompañado de máximo seis (6) palabras claves que describan los tópicos, áreas o temas más importantes del trabajo desarrollado.

El desarrollo del tema usualmente contempla secciones de Introducción, Revisión, Métodos, Materiales, Experimentos, Resultados, Discusión, Conclusiones, Agradecimientos,

Referencias y Apéndices. Cada sección va numerada con notación arábiga, con subsecciones numeradas en nivel de profundidad (i.e. 1., 1.1, 1.2, 1.3, 2., 2.1, etc.). Los resultados se pueden presentar en Tablas o Figuras, referenciadas desde el texto.

Las figuras van numeradas en orden de aparición en el texto del artículo con números arábigos, precedidos por la palabra “Figura” y en negrilla. Todas las figuras deben contar con una leyenda explicativa. Las imágenes deben tener una resolución mínima de 200 píxeles por pulgada. Las figuras deben ser en lo posible en tonalidades de gris, aunque también se aceptan en color si es necesario para el entendimiento del artículo. Las ecuaciones dentro de un artículo deben estar numeradas en orden de aparición. El número de ecuación va entre paréntesis y ajustado al lado derecho de la página mientras que la ecuación va centrada. Las tablas van numeradas en forma consecutiva en la medida que van apareciendo en el texto del artículo. Las tablas van encabezadas por la palabra “Tabla” y un número romano consecutivo, todo en negrilla. Todas las referencias bibliográficas deben estar citadas dentro del texto. Las referencias aparecen al final del artículo, después de las conclusiones. Las referencias deben listarse de manera numerada en el orden de aparición en el texto del artículo, con números arábigos entre corchetes cuadrados. El formato de presentación de referencias que acepta la Revista es el IEEE.

El artículo debe finalizar con un espacio dedicado a las biografías de los autores. En ellas deben aparecer: nombre completo, lugar de origen, título de pregrado, universidad, ciudad, país, títulos de postgrado, universidad, ciudad, país, experiencia profesional (cargo, institución, años de servicio) y datos de contacto.

4. PROCEDIMIENTO PARA EL ENVÍO DE ARTÍCULOS PARA PUBLICACIÓN

Se describe el procedimiento que deben seguir los autores para enviar artículos para publicación en la Revista INGENIERÍA.

Los autores deben enviar los siguientes documentos por correo electrónico (revista_ing@udistrital.edu.co) dirigidos al Editor de la Revista:

- Carta de presentación y originalidad del artículo: El formato de la presente carta puede descargarse en el siguiente link: <http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Formatos>. La finalidad de la carta es certificar que el artículo es inédito, que no está sometido a otra publicación y que se aceptan las normas de publicación de la Revista.
- El artículo en formato digital utilizando la plantilla y las normas de presentación que se pueden consultar en el siguiente link (instrucciones a los autores): <http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Instructivos>. Se deben enviar dos (2) copias del artículo, una de ellas sin los nombres ni biografías de los autores.
- Los autores relacionaran en el correo de postulación del artículo dos posible evaluadores que cumplan las siguientes condiciones:
 - Debe tener al menos título de maestría
 - Pertener a una institución diferente a la de los autores
 - Haber publicaciones en los últimos tres años

Nota: Enviar los datos de los contacto de los evaluadores (nombre completo, institución y correo electrónico) en el correo.

- En caso de ser aceptado, los autores del artículo se comprometen a diligenciar el formato de cesión de derechos para publicación. Se puede descargar en el siguiente link: <http://revistas.udistrital.edu.co/ojs/index.php/reving/pages/view/Formatos>