Decision Support Systems

Research in Computing Science

Research in Computing Science es una publicación trimestral, de circulación internacional, editada por el Centro de Investigación en Computación del IPN, para dar a conocer los avances de investigación científica y desarrollo tecnológico de la comunidad científica internacional. **Volumen 147, No. 3,** marzo 2018. Tiraje: 500 ejemplares. *Certificado de Reserva de Derechos al Uso Exclusivo del Título* No.: 04-2005-121611550100-102, expedido por el Instituto Nacional de Derecho de Autor. *Certificado de Licitud de Título* No. 12897, *Certificado de licitud de Contenido* No. 10470, expedidos por la Comisión Calificadora de Publicaciones y Revistas Ilustradas. El contenido de los artículos es responsabilidad exclusiva de sus respectivos autores. Queda prohibida la reproducción total o parcial, por cualquier medio, sin el permiso expreso del editor, excepto para uso personal o de estudio haciendo cita explícita en la primera página de cada documento. Impreso en la Ciudad de México, en los Talleres Gráficos del IPN – Dirección de Publicaciones, Tres Guerras 27, Centro Histórico, México, D.F. Distribuida por el Centro de Investigación en Computación, Av. Juan de Dios Bátiz S/N, Esq. Av. Miguel Othón de Mendizábal, Col. Nueva Industrial Vallejo, C.P. 07738, México, D.F. Tel. 57 29 60 00, ext. 56571.

Editor responsable: *Grigori Sidorov, RFC SIGR651028L69*

Research in Computing Science is published by the Center for Computing Research of IPN. **Volume 147**, **No. 3**, March 2018. Printing 500. The authors are responsible for the contents of their articles.

Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

Decision Support Systems

Cuauhtémoc Sánchez-Ramírez
Giner Alor-Hernández
Jorge Luis García-Alcaraz (eds.)

SEP
SECRETARÍA DE EDUCACIÓN PÚBLICA

Instituto Politécnico Nacional
"La Técnica al Servicio de la Patria"

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2018

ISSN: 1870-4069

Instituto Politécnico Nacional (IPN)
Centro de Investigación en Computación (CIC)
Av. Juan de Dios Bátiz s/n esq. M. Othón de Mendizábal
Unidad Profesional "Adolfo López Mateos", Zacatenco
07738, México D.F., México

http://www.rcs.cic.ipn.mx
http://www.ipn.mx
http://www.cic.ipn.mx

Indexed in LATINDEX, DBLP and Periodica

Printing: 500

Printed in Mexico

ISSN 1870-4069

Editorial

The uncertainty present in industrial processes requires the use of tools, methodologies and techniques that allowed better making decisions to managers, for this reason the use of systems integrated into business technologies are usually called Decision Support Systems (DSS). In this volume eight papers are presented that were carefully selected of 12 submissions about the use of different techniques for designing and developing Decision Support System (DSS) in industrial contexts. These papers were evaluated by an editorial board integrated for reviewers with international prestige in this area. The papers were selected by considering the originality, scientific contribution to the field, soundness and technical quality of the papers.

Some used techniques about DSS in industrial contexts presented in this volume are: 1) A conceptual model for the design of a renewable energy supply chain from biomass; 2) Design and Implementation of a Data Warehouse to Support Decision-making in a Health Environment; 3) Agile Dimensional Model for a Data Warehouse Implementation in a Software Developer Company; 4) Multiobjective model to reduce logistics costs and CO_2 emissions in goods distribution; 5) Multi-Agent Model for Urban Goods Distribution; 6) A CBIR System for the Recognition of Agricultural Machinery; 7) The Role of ICT in the Supply Chain of Ciudad Juarez Industrial Sector; 8) Ontology - based Operational Risk Identification in 3PL, among other themes.

The editors would like to express their gratitude to the reviewers who kindly contributed to the evaluation of papers at all stages of the editing process. They equally thank the Editor-in-Chief, Prof. Grigori Sidorov, for the opportunity offered to edit this special issue and for providing his valuable comments to improve the selection of research works. Guest editors are grateful to the National Technological of Mexico for supporting this work. This book was also sponsored by the National Council of Science and Technology (CONACYT) as part of the project named Thematic Network in Industrial Process Optimization, as well as by the Public Education Secretary (SEP) through

Cuauhtémoc Sánchez-Ramírez (Instituto Tecnológico de Orizaba, Mexico)
Giner Alor-Hernández (Instituto Tecnológico de Orizaba, Mexico)
Jorge Luis García-Alcaraz (Universidad Autónoma de Ciudad Juárez, Mexico)
Guest Editors
March 2018

ISSN 1870-4069

Table of Contents

Page

ISSN 1870-4069

A CBIR System for the Recognition of Agricultural Machinery

Rodolfo Rojas Ruiz[1], Lisbeth Rodríguez-Mazahua[1], Asdrúbal López-Chau[2], Silvestre Gustavo Peláez-Camarena[1], María Antonieta Abud-Figueroa[1], Isaac Machorro-Cano[1]

[1] División de Estudios de Posgrado e Investigación, Instituto Tecnológico de Orizaba, Orizaba, Veracruz, México

[2] Universidad Autónoma del Estado de México, Centro Universitario UAEM Zumpango, Zumpango, Estado de México, México

Rudy_199191@hotmail.com, lrodriguez@itorizaba.edu.mx, imachorro@gmail.com, {gpelaez, mabud}@ito-depi.edu.mx, alchau@uaemex.mx

Abstract. Content-Based Image Retrieval (CBIR) is a technique to obtain images based on visual characteristics like color, texture and shape. In this article, an analysis of different CBIR systems is presented; also an inventory system that allows content-based queries is proposed to help the employees of a company that sells agricultural machinery and spare parts, called AGROMAQ. This system identifies different products, all with the aim of reducing customer service time, since an expert is currently needed to identify them, which increases attention time not only in hours but also in days.

Keywords: CBIR system, BoofCV, SURF descriptor.

1 Introduction

Multimedia data provide a direct representation of the physical world in the digital format and include photographs, X-ray images, sound, and video [1]. A multimedia database is a collection of multimedia data that provides: 1) Content-based access; 2) Generation of knowledge; 3) Handling of large volumes of data, and 4) Good response times [2].

This article presents an inventory system for the company AGROMAQ. The system uses CBIR to perform the search for spare parts and machinery based on an image, everything with the aim of improving customer service times and reducing the dependence on an expert.

The descriptor SURF (Speeded Up Robust Features) was used as a method for the extraction of characteristics of the pieces and their subsequent automatic recognition. SURF [3] is one of the most widely used algorithms in the literature for extraction of invariant points of interest. The extraction of the points is done by first detecting the possible points of interest and their location within the image. Subsequently, the neighborhood of the point of interest is represented as a feature vector (by default with a size of 64, although it is possible to increase it). The SURF algorithm uses a basic approximation of the Hessian matrix to reduce computation time. The Hessian matrix is used because of its good relationship between accuracy and time cost. The main feature of SURF points of interest is repeatability, if the point is considered reliable, the detector will find the same point under different perspectives (different scale, orientation, among others).

This article is made up of five sections. Section 2 in this paper is an analysis of different articles related to CBIR. Section 3 describes the different phases of the CBIR system. The evaluation of the system is discussed in Section 4. Finally, the conclusions and future work are presented in Section 5.

2 State of the Art

Navel et al. [4] proposed an enhanced SURF for CBIR systems, the enhanced was made by extracting Hu moments and eigen values from the image thus forming a 88 dimension enhanced SURF descriptor. The experiments with this enhanced version of SURF proved that this version is better than the original SURF. In contrast, Kommineni and Chava [5], developed a CBIR system that uses the SIFT (Scale-Invariable Feature Transform) descriptor in conjunction with SPM (Linear Spatial Pyramid Matching), which improves object recognition and results in Sparse + SIFT.

Table 1. Comparison of the related works.

Article	Evaluation	Descriptor name
[3]	Enhanced SURF descriptor is better than SURF in precision (increased in 9.2%) and recall (increased in 3%).	Enhanced SURF
[4]	Sparse+SIFT descriptor proved to be more accurate than the SIFT descriptor.	Sparse+SIFT
[5]	The CBIR system has a precision in the range between 69.51% and 91.19%.	SURF
[6]	The CBIR system has a precision of the 94.2% with the use of geometric constraints.	SIFT
[7]	SURF with standard deviation works better to find images in a database with sketches.	SURF
[8]	The CBIR system has a precision of the 73.17% using SIFT as descriptor.	SIFT
[9]	The precision of the system was 80% for MRI, 75% for Iris and 90% for Bones.	SURF
[10]	Dynamic delimitation is more effective and robust than static delimitation.	TOP-SURF JCD (Joint Composite Descriptor).
[11]	LSH (*Locality-sensitive hashing*) was the fastest technique.	LSH
[12]	The proposed method has an average precision of the 78.97%	SURF

Ronald et al. [6] constructed a CBIR system, Plantcyclopedia, that can identify plants given by the user using SURF, is divided into three main steps: 1) Retrieve the image information; 2) Define the orientation of the image, and 3) Obtain feature points from both images using K-Means as a classification technique. Lee et al. [7] presented a CBIR system for a database of tattoo images based on SIFT. The goal of the system is to obtain visually similar images to the query image as well as determining some extra feature to the shape and design of the tattoo, such as the area of the body where it is located and class tags of tattoos.

Pradnya and Pravin [8] developed a SBIR (Sketch Based Image Retrieval) system, this system uses SURF with a standard deviation to recover the images from the database. Whereas SIFT was used in a system for detection and pairing of drug pills presented in [9], the system used as the basis the detection of the pills depending on their shape, stamp / figure, color, color histogram and Hu moments. Kaur and Jindal [10] proposed a feature extraction on medical images using SURF under the Open CV platform. The steps of the system are five: 1) Load test images; 2) Extract key points, contours and textured data; 3) Store the features in a vector; 4) Display images highlighting key points, and 5) Compute performance.

Avi et al. [11] presented a two-step method for a CBIR system in large multimedia databases, in order to improve both the effectiveness and efficiency of the traditional CBIR by exploring secondary media, such as pre-filtered text or images. The two-step recovery method at the beginning classifies by a secondary means and then performs CBIR only on the best ranked K-objects. The proposal of an architecture for efficient image recovery, as well as Web analysis from the point of view of images, using Big Data technologies, is the main theme on [12]. The phases of the proposed architecture are: 1) Signature of the input image; 2) Search with binary trees; 3) Distributed calculation of distances, and 4) Ordination of similar images.

Alexandra et al. [13], proposed a method for CBIR and classification, the method used SURF combined with Bag-of-Visual-Words (BoVW). This combination has a good retrieval and classification results compared to other methods. SURF has four major steps: 1) Integral image, 2) Fast Hessian, 3) SURF descriptor, and 4) Salient features. BoVW consists of three main steps: 1) Automatically extract the interest points and descriptor from the images, 2) Quantize the keypoints and descriptors to form the visual dictionary, and 3) Find the occurrences of each visual words in the image in order to build the BoW histogram.

In conclusion, CBIR systems are important because they facilitate the task of identifying objects to people who are not experts in the field, and are not limited to a single area such as medicine, botany, forensic medicine, among others. Finally, it is possible to say that SIFT and SURF have good results (precision) in the CBIR systems as shown in Table 1. Nevertheless, Section 4 will show that the time and recall are better in SURF than SIFT, thus, SURF will be used in the CBIR system of the company AGROMAQ.

3 CBIR System

Figure 1, shows the process followed by the system for the extraction and storage of the descriptors each time an image is recorded and thus saving time in the search.

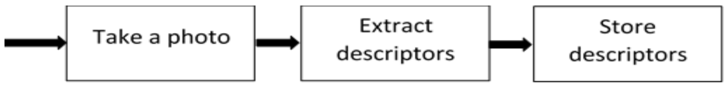


Fig. 1. Process during the recording of images.

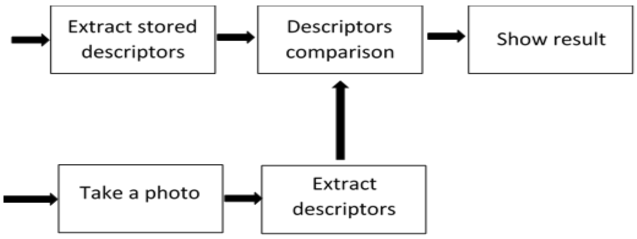


Fig. 2. CBIR process.

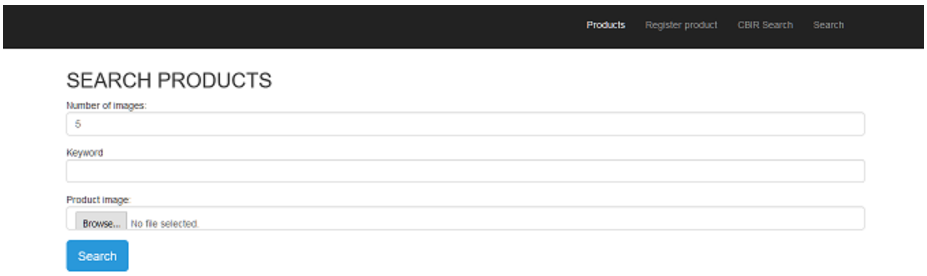


Fig. 3. Interface of the content based query.

During the registration of a product the user takes one or several photographs of the product to be registered and uploads them to the system. Once the product is registered, the system extracts the points of interest of each of the images and the descriptors using SURF, finally, these descriptors are stored in the form of objects for their future use.

In Figure 2, on the other hand, it is represented the process that the system follows for the image-query and the images in the database during the search process. The user chooses to write a search field to help reduce the sample of products to be compared during the process. If this is done, the system will search for images with those fields in their description and will return those that fit. After doing this, it extracts the descriptors stored in objects of each of these images, in case no search field has been written, the system will take the complete sample of all the registered images.

Likewise, the user must enter a photograph of the product to be searched to which the points of interest will be extracted and the descriptors will be calculated again with SURF. Once the descriptors of both the images in the database and the query image are obtained, the system proceeds to compare the query image with each of the images in the database to obtain the points where the pairing is correct, once this process is finished, the images are sorted with respect to the number of correct pairings and the results are displayed.

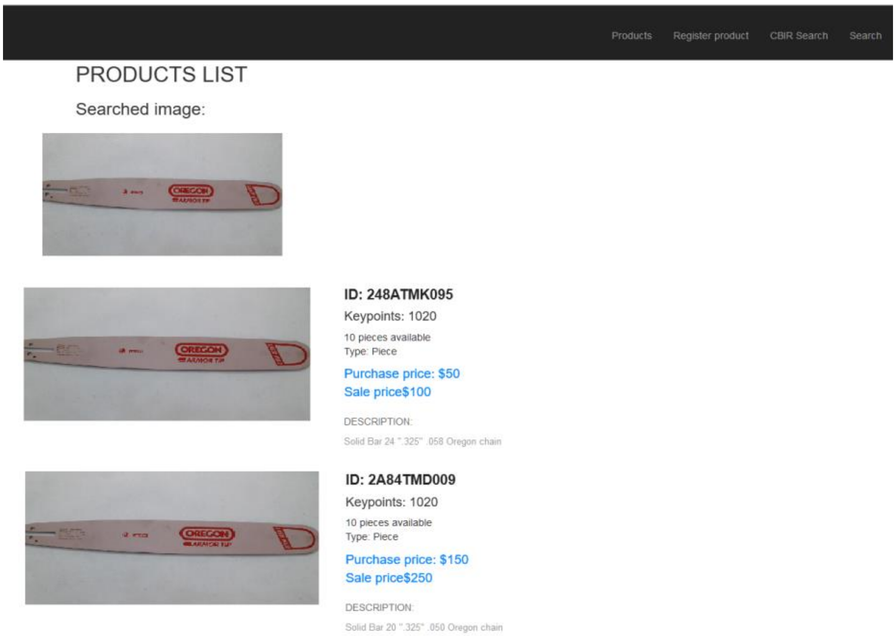


Fig. 4. Interface of the result of the content based query.

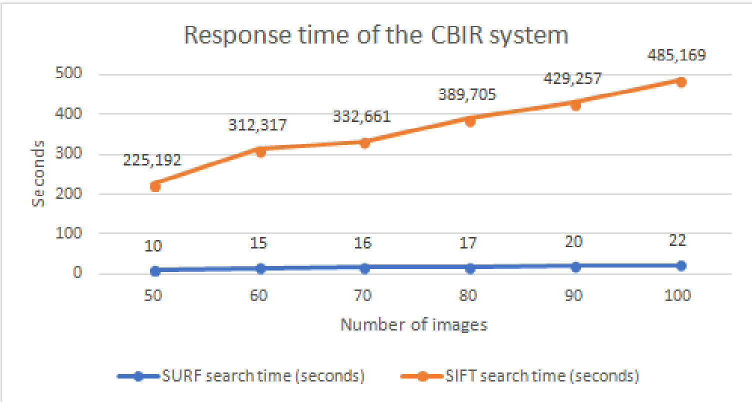
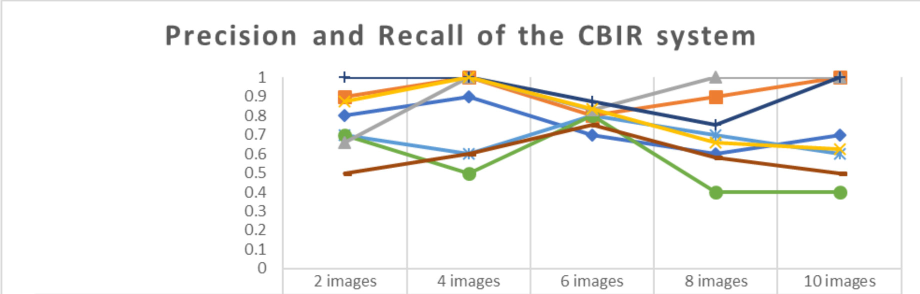


Fig. 5. Comparative diagram of efficiency using SIFT and SURF.

The CBIR system was implemented using Java with the JavaServerFaces framework. The multimedia database was implemented using PostgreSQL. The search interface in Figure 3 shows two fields to proceed, which are the number of images most similar to the given image, and the search image (query). Once the search is done, the system displays the results in list form, detailing the product or products that are similar to the given image as shown in Figure 4.



	2 images	4 images	6 images	8 images	10 images
Machinery precision (SURF)	0.8	0.9	0.7	0.6	0.7
Machinery precision (SIFT)	0.7	0.6	0.8	0.7	0.6
Piece precision (SURF)	0.9	1	0.8	0.9	1
Piece precision (SIFT)	0.7	0.5	0.8	0.4	0.4
Machinery recall (SURF)	0.66	1	0.83	1	1
Machinery recall (SIFT)	1	1	0.875	0.75	1
Piece recall (SURF)	0.875	1	0.833	0.66	0.625
Piece recall (SIFT)	0.5	0.6	0.75	0.58	0.5

Fig. 6. Comparative diagram of effectiveness between SIFT and SURF.

4 Experiments

There are two main parameters used to measure the performance of a CBIR system: efficiency and effectiveness. Efficiency is mainly concerned with the response time of the system, while effectiveness is related to the system's ability to retrieve relevant items and discard irrelevant items [13]. In order to evaluate the performance of the CBIR system presented in this paper, the follow experiments were made with the descriptor SURF from BoofCV v0.26 and the descriptor SIFT from OpenCV 2.4.13.

4.1 Efficiency of the CBIR System

A performance evaluation was realized on the response time of the system, the content-based query of the figures 3 and 4 was executed in different cases which varied from each other by the number of images in the database, ranging from 50 images to 100 images increased 10 images in each case both for the SIFT and SURF descriptors. In all cases the number of images retrieved was 5.

Figure 5 shows clearly that the CBIR system applying the SURF descriptor has a better efficiency than when it implements SIFT.

4.2 Effectiveness of the CBIR System

Based on the state of the art presented in Section 2, recall and precision are the most used measures to evaluate the effectiveness of CBIR systems. Recall measures the ability to retrieve relevant items, and precision measures the ability to reject irrelevant items. A good system should have both high recall and precision values [3, 14]. Therefore, during the tests performed in the system where the evaluation was realized to two different groups of images (Pieces and Machinery) in 10 iterations, applied to the 100 images in the database, the results that were obtained by equations (1) and (2) are shown in Figure 6. Five cases were considered; 2, 4, 6, 8 and 10 queries executed and in all cases the total number of images retrieved was 5.

$$\textbf{Recall} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in database}}, \tag{1}$$

$$\textbf{Precision} = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}. \tag{2}$$

Figure 6, shows a comparison between SIFT and SURF and as it can be seen the effectiveness of the CBIR system using SURF is better most of the times taking into account the precision and the recall.

5 Conclusion and Future Work

CBIR systems are important at the commercial level because they make easier to identify objects, spare parts and machinery in this case, to people who are not experts in the field, as well as the fact that they no longer depend so much on experts in the area, with the aim to make customer service times shorter.

In this article, a CBIR system for the recognition of agricultural machinery and spare parts was showed, also the efficiency and effectiveness of the system was evaluated through the results of the experiments realized with SIFT and SURF descriptors as well as a brief comparison between them, with those results it was concluded that in the case of this article SURF is faster than SIFT and has better performance in precision and recall.

The performance of the CBIR process will be tried to improve by implementing a system that allows an environment in which it is possible to have a control of the brightness during the shooting of the products.

Acknowledgments. The authors are very grateful to the National Council of Science and Technology (CONACYT) for supporting this work. Also, this research paper was sponsored by National Technological of Mexico, as well as by the Public Education Secretary (SEP) through PRODEP.

References

1. Phillip, K.C.: Multimedia Information Storage and Retrieval. Techniques and Technologies, IGI Publishing, Hershey (2008)
2. Assent, I.: Efficient adaptive retrieval and mining in large multimedia databases in Dissertation. Fakultät für Mathematik, Informatik und Naturwissenschaften, RWTH Aachen (2008)
3. Romero, M. et al.: Comparativa de detectores de características visuales y su aplicación al SLAM in Comparativa de detectores de características visuales y su aplicación al SLAM. Cáceres, pp. 55−62 (2009)

4. Navel, G. et al.: Content Based Image Retrieval using Enhanced SURF in Computer Vision, Pattern Recognition, Image Processing and Graphics (2015)
5. Kommineni, D., Chava, S.: CBIR Based On Linear SPM. Using SIFT Sparse Codes in IEEE International Conference on Computational Intelligence and Computing Research (2016)
6. Ronald, A. et al.: Plant Encyclopedia with Image Matching on a Web-Based Application. In International Conference on Information Management and Technology, Indonesia, pp. 89−92 (2016)
7. Lee, J. et al.: Image Retrieval in Forensics: Tattoo Image Database Application in Multimedia in Forensics. Security and Intelligence, 19(1), pp. 2−11 (2012)
8. Content Based Image Retrieval by Using Sketches: Information Communication and Embedded Systems (2014)
9. Lee, Y. et al.: Pill-ID: Matching and retrieval of drug pill images. In Pattern Recognition Letters, 33, pp. 904−910 (2012)
10. Kaur, B., Jindal, S.: An implementation of Feature Extraction over medical images on OPEN CV Environment. In Devices, Circuits and Communications (2014)
11. Avi, A. et al.: Dynamic two-stage image retrieval from large multimedia databases. In Information Processing and Management, 49, pp. 274−285 (2013)
12. Sergui, R. et al.: Tecnologías big data para análisis y recuperación de imágenes web en el profesional de la información, 23(6), pp. 567−574 (2014)
13. Alexandra, A. et al.: Content Based Image Retrieval and Classification Using Speeded-Up Robust Features (SURF) and Grouped Bag-of-Visual-Words (GBoVW). In International Conference on Technology, Informatics, Management, Engineering & Environment, Indonesia, pp. 77–82 (2013)
14. Lu, G.: Multimedia Database Management Systems. Artech House Computing Library (1999)

ISSN 1870-4069

A Conceptual Model for the Design of a Renewable Energy Supply Chain from Biomass

Ramos-Hernández Rocío[1], Sánchez-Ramírez Cuauhtémoc[1], Alor-Hernández Giner[1], González-Huerta Magno-Ángel[1], Jiménez-Nieto Yara-Anahi[2], Rodríguez-Parada Adolfo[2]

[1] Division of research and postgraduate Studies, Instituto Tecnológico de Orizaba, Orizaba, Veracruz, Mexico

[2] Faculty of accounting and administration, Universidad Veracruzana campus Ixtaczoquitlán, Veracruz, Mexico

Abstract. Renewable energy (RE) is one of the areas where major research and development efforts are currently developing. Demand for electrical energy is increasing because of population growth and accelerated industrialization, particularly in developing countries like India and China. Traditional forms of energy generation produce large amounts of CO^2, which adversely affects the environment. Renewable energies include solar, wind, geothermal, hydropower and biomass (BM), the latter being the option with the greatest potential for electricity generation in Mexico [1], which ranks 20th in the world in producing energy from BM. However, identifying regionally important types of BM to evaluate their energy potential, localities of conversion and demand, besides integration of the conventional electrical system, represents technological challenges in administration and control. This generates high logistics costs and the need for good supply chain planning (CS). The efficient management of CS of bioenergy and logistics for key processes (procurement, production and distribution), represents key issues [2]. According to some authors, a new trend has arisen; where it is not enough to compete between organizations; instead, the competition shifts towards the terms of CS of RE and traditional forms of energy generation. This paper presents the conceptual development of a bioenergy supply chain in Mexico, using system dynamics as a methodology for the design and evaluation of the supply chain, since this is a suitable approach to model and study the interaction between variables in a complex system by applying feedback loops [3].

Keywords: system dynamics, renewable energy, biomass, supply chain.

1 Introduction

The fossil energies (FE) are the most commonly used energies worldwide; however, they have two major disadvantages: 1) they use non-renewable sources and, 2) they are aggressive to the environment; because of this, studies for the implementation of renewable energies have increased in the last decades. In addition, many countries are seeking to replace FE with renewable energies (RE). An RE is described as "*Energy derived from a natural process that is constantly replenished*"[4]. One of the leading countries on the subject is Spain, which has evolved as one of the countries with the best generation of electricity in areas where solar and wind energy are most feasible [5]. Conolly et al [6] carried out a study in which they proposed to achieve the generation of electricity through RE to fulfill 100% of energy demnd by the European Union. According to Gold and Seauring [2], within the concept of RE technologies, Bioenergy will play a decisive role in the following decades and under favorable conditions and based on an intelligent design can be applied for electricity generation. Bioenergy is the energy obtained from biomass, which is defined as "*Biodegradable part of products, waste and residues from agriculture, forestry and related industries, as well as the biodegradable fraction of industrial and municipal waste*" [7].

Although the country is highly rich in RE sources, Mexico´s potential for generating this type of energy has not been fully exploited. However, the diversification of primary energies represents an opportunity for the strengthening of national energy security, as the biomass industry has been expanding rapidly in recent years. The Official Gazette of the Federation [8, 9] lists some of the technological challenges faced by efforts to expand the country´s capacity to generate energy by this strategy, including, 1) The viability of its incorporation into the national electricity system, 2) The different scenarios for the participation in generating electricity from (BM), 3) The evaluation of technological options within the production scenarios, 4) The implications of energy efficiency in production, transformation and generation of energy from BM and, 5) The costs of the logistic processes of supply, generation and marketing of bioenergy (supply chain).

The main objective of a bioenergy supply chain is to produce energy in an economically viable way. However, uncertainty propagates in spatial and temporal dimensions through it, which significantly affects the performance of the system [10]. This makes the design and development of the supply chain one of the most critical points. The main contribution of this research is, therefore, a conceptual model for the design of the biomass supply chain for the generation of electricity, considering the key logistics processes (procurement, production and distribution). Similarly, evaluation of the model uses system dynamics (SD) because this is an approach to model and study the interaction between variables in a complex system by applying feedback loops [3].

2 Tools for the Study of Renewable Energies

Although studies on renewable energies have increased in the last couple of decades, they do not yet allow a broad idea of resources to be analyzed but partially.

However, Engelken et al. [11]; Gatzert and Kosub [12]; Yudi and Sofri [13]; Goh et al. [14]; Lund [15]; Hitzeroth and Megerle [3] and Lokey [16] offer a rather detailed analysis of the different risks faced when designing, implementing and managing a renewable energy plant. Aslani and Mohaghar [17] and Andersen et al. [1] propose a

business model for the implementation of some type of renewable energy, such as the installation of stations to recharge electric cars. On the other hand, Loong Lam et al. [18]; López et al. [19]; Mikatia et al. [20]; Rylatt et al. [21], among others, use different techniques, such as P-Graph, Road map, MatLab and agent-based simulation, to solve different aspects of renewable energies, such as the power needed to implement a photovoltaic plant. Renewable energies possess many advantages, but they also present some major challenges to the supply system. One of the most demanding features is the uncontrolable variability of the source, since it is not possible to guarantee environmental conditions in advance and, consequently, neither the supply.
Although there are different tools that handle uncertainty using stochastic programming, optimization, fuzzy programming or some other optimization technique under uncertainty [22], in this project, systems dynamics will be the methodology to use, since it is an approach to model and study the interaction between variables in a complex system by applying feedback loops [23].

2.1 Supply Chains in Renewable Energies and Systems Dynamics

The Supply Chain (SC) is a network of interdependent and connected organizations that work in a coordinated way to control, manage and improve the flow of materials and information from raw material suppliers to customers [24].

Rendon [25] and Ramos [26] have described some of the techniques and tools used in the design and evaluation of SC. Rendón analyzed the SC of biofuel generation through ethanol and Ramos analyzed the SC of the creation of a new product generated from a waste. D 'Amore et al. [27] investigated a supply chain considering cultivation of bio mass, transport, conversion to bio ethanol or bioelectricity, distribution and its final use as bio fuel and electric cars. Azadeh and Vafa [28] presented a hybrid systems dynamics/mathematical programming model to design and plan a bio-diesel supply chain. The performance of the bioenergy supply chain depends highly on good coordination between links. A whole range of strategies can be adopted for the efficient design and evaluation of the SC to ensure that these interrelationships work; the methodology used in this project is outlined below.

3 Methodology to Design a Bioenergy Supply Chain

Figure 2 represents the methodology for setting up the conceptual model of the design and evaluation of a bioenergy supply chain. The methodology consists of 6 main steps: 1) Search of variables; 2) Definition of relations between the variables 3) Implementation of the diagram to blocks; 4) Elaboration of the causal diagram; 5) Evaluation of causal relationships and, 6) Final conceptual model. These steps are described below.

1. ***Variable search***. This is carried out through a systematic research of the state of the art concerning the key words: biomass, renewable energy, supply chain, supply chain design and supply chain evaluation.

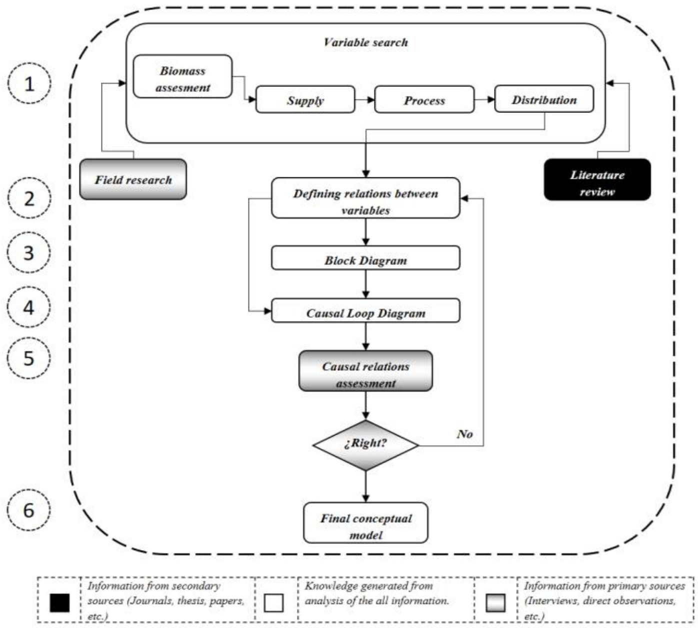


Fig. 1. Methodology for the design and evaluation of a bioenergy supply chain.

2. ***Definition of relations between variables***. Each variable identifies the existing primary relationships with other variables in addition to the link in which they are within the supply chain.
3. ***Block Diagram.*** With the primary relations already identified, the block diagram is elaborated.
4. ***Causal diagram***. Once the main relationships have been identified, the most complex relationships between two or more variables, as well as the manner in which they impact each other (positively or negatively) are set. The feedback loops and the type to which they belong (Balance or Reinforcing), are also indicated.
5. ***Evaluation of causal relationships***. Once the causal diagram is finished, an evaluation of the causal relationships, the type of impact and the interconnected variables are representative of the system to be evaluated.
6. ***Final conceptual model.*** The final conceptual model is the causal diagram with the causal relationships already evaluated. This diagram will be base for the simulation of the supply chain for its evaluation and generation of policies and strategies that allow its development.

This article presents the steps for the development of this methodology, which culminates in the elaboration of the causal diagram that gives rise to the final conceptual model. The steps are described in the next section.

4 Results of the Application of the Methodology for the Design of a Bioenergy Supply Chain

This section presents the steps to follow the methodology for the conceptual model of design and evaluation of a general bioenergy supply chain.

4.1 Steps 1 and 2 - Search for Variables and Definition of Relationships Between Variables

Supply, production and distribution are the three main parts of a supply chain [29], however, for the design of the bio-energy supply chain, it is necessary to add an extra link for the evaluation of biomass potential. The importance of these four parts is described below.

Biomass potential. The estimation of the potential energy of the biomass residue requires an evaluation of these residues. Authors such as Paiano and Lagioia [30] recommend a qualitative evaluation to identify the crop that best adapts to the needs of the study (region, cultivation area and harvesting period, among others).

Procurement. The purpose of procurement is control of supplies in order to meet the needs of operational processes. [31].

Production. The production process converts raw materials into finished products, and in the supply chain of bioenergy generation the process becomes even more complex when using anaerobic digestion, since it must consider variables that are mostly unstable, such as Ph, temperature and bacterial growth, among others [32].

Distribution. It ensures that these final products reach the consumer through a network of distributors; however, the final product being the electrical energy, the considerations to take are more. These considerations depend on the analysis for incorporation into the national electricity system and their respective policies [33].

Due to the importance of the links mentioned above, the variables to be searched are focused on each one of them and in locating which link they are in. Table 1 lists the variables that are in the block diagram and that will later integrate the causal diagram.

4.2 Step 3 - Block Diagram of the Design of a Bioenergy Supply Chain

This section analyzes the interaction between the links in the supply chain and the assessment of the availability of biomass. These elements are part of the conceptual model proposed for the design of the bioenergy supply chain and will be analyzed by means of a diagram of blocks. Figure 1 shows the block diagram of the operation of a general bioenergy supply chain. The interactions work as described below:

Table 1. Variables in the links of a bioenergy supply chain.

Authors	Variables	Authors	Variables	Authors	
11, 12	Crop area	14, 27	Process yield	30,33	Supplier's storage
	Hectare average		Harvest waste		Supplier's storage capacity
14, 27, 33	Raw material		Processing waste		Distance
	Herbaceous		Available waste		Supplier's transport
	Arboreal		Unusable waste		Transport cost
	Forest	32	Initial transport		Supplier's transport capacity
	Growth rate		Storage cost		
	Decreasing rate	30	Biomass as raw material	34	Biomass conversion plant's storage
	Yield per hectare	28	Biomass allocation	3	Conversion process capacity
	Harvest		Type	6,13	Conversion process to electricity
	Process		Master program		
15, 16, 18, 28	Costs	1, 7, 33	Standards	29, 32	Distribution
	Demand		Substrate		Anaerobic reactor

1. A biomass availability assessment is performed, based on the type of raw material (RM), its area of sowing and the average yield considering the growth rates and decrease of the type of crop.
2. The available biomass is transported to a supplying company (Supply), where the reduction obtained in previous steps is converted to RM for power generation plants.
3. The BM received at the electric power plant (Production) must pass through a process of conversion (Biomass to energy).
4. Finally, the energy generated must be distributed through the lines of the national electricity system.

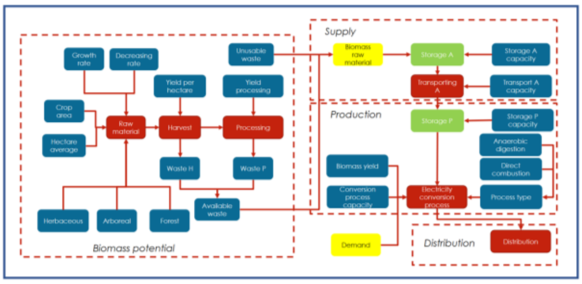


Fig. 2. Block Diagram of the Bioenergy Supply Chain.

4.3 Step 4 and 5 - Causal Diagram and Evaluation of Causal Relationships

Once the block diagram is constructed, these primary relationships are plotted in a causal diagram, which is likewise a graphical representation of the interactions between the variables of a system. However, more complex relationships can be represented between them, as well as how another variable is affected either positively or negatively, in addition to feedback loops.

The result of having captured the primary relationships in a causal diagram, together with the identification of the complex relationships and feedback loops generated is shown in Figure 3. As can be seen, several complex relationships are generated along with 3 balancing loops and 2 of reinforcement, which will be explained later.

The feedback loops generated in the causal diagram are described below.

B1. *Supplier's storage* negatively impacts the *Back orders* since, if there is more biomass in the supplier's warehouse, then the number of pending orders may decrease; on the other hand, if there is not enough biomass, then the number of outstanding orders will continue to grow. *Biomass ordered* is affected by back orders in a positive way.

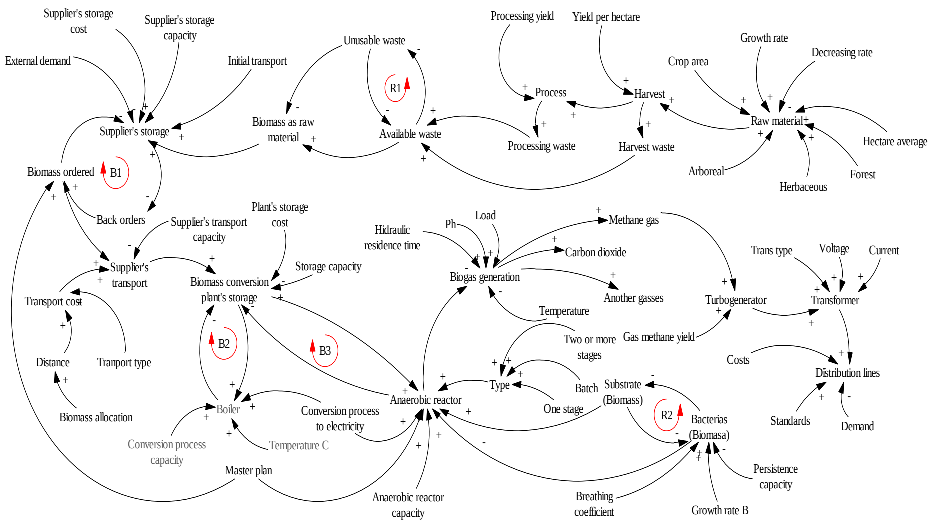


Fig. 3. Causal diagram of a bioenergy supply chain.

B2. *Biomass conversion plant's storage* positively impacts the *Boiler* variable because, if the storage of biomass is at a good level, it will be more feasible to carry out the conversion process through a direct combustion boiler; on the other hand, if there is not have enough of biomass, then the conversion process will be hindered. On the other hand, the variable *Boiler* negatively impacts to the *Biomass conversion plant's storage* if indeed greater production goes together with smaller volume in inventory.

B3. The relationship between the variables *Biomass conversion plant's storage* and *Anaerobic reactor* lies in the inventory levels, since the higher the available biomass, the greater the feasibility of carrying out the conversion process. Similarly, the higher biomass demands for the reactor, the lower amount of biomass available for processing.

R1. The variables Unusable waste and Available waste generate a loop where, the greater the amount of available waste, the smaller the amount of waste that is unusable and vice versa.

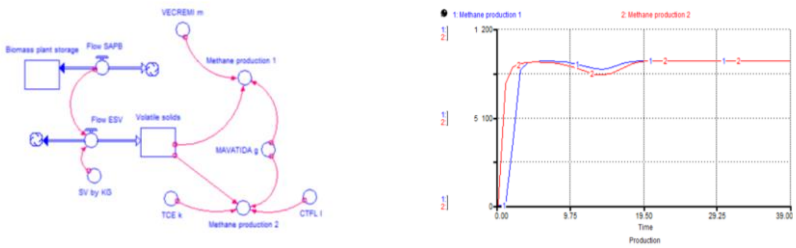
R2. The variables Substrate (Biomass) and Bacteria (Bio-mass) reinforce the idea of having 100% organic matter inside the reactor.

After performing the causal diagram, validation tests of causal relations are performed. From these tests, along with the experts in the matter we conclude that the causal diagram is representative of the system.

4.4 Step 6 of the Methodology Followed - Final Conceptual Model

Once the causal relationships of the causal diagram have been evaluated and its representativity of the system is verified, it is taken as the final conceptual model. This model serves as reference so that the system can be captured in a Forrester diagram and then, simulated. As an example, a fraction of the causal diagram of the SC of bioenergy, which represents the generation of biogas, is presented. Figure 4b represents the Forrester diagram and Figure 4c the results of simulation.

Although this step is terminated in the present methodology, there are still steps to follow within the methodology of system dynamics for the design and evaluation of a SC of bioenergy, as indicated in the section of conclusions and future work.



a. Diagrama de Forrester b .Simulación del comportamiento de biogas

Fig. 4. Generation of biogas in a bioenergy supply chain.

4 Conclusions and Future Work

Using the system dynamics and causal diagram approach, this paper proposes a model for designing and evaluating a biomass power supply chain based on the regional needs and biomass available in the geographic region. We verified that system dynamics is an adequate approach for this type of studies, since it is a tool that allows to integrally model the different interactions between the links of the supply chain and the biomass evaluation link available. Likewise, the model provides the basis for the construction of simulation models for the design and evaluation of supply chains, not only of different types of biomass, but also of different types of renewable energies. As a future work, the proposed conceptual model will be applied in a case study following the steps described in the methodology section, with the aim of validating the model in its field of application and thus obtaining feedback from the system.

References

1. Andersen, P. H., et al.: Integrating private transport into renewable energy policy: The strategy of creating intelligent recharging grids for electric vehicles. Energy Policy (2009)
2. Gold, S., Seuring, S.: Supply chain and logistics issues of bio-energy production. Journal of Cleaner Production, pp. 32–42 (2010)
3. Hitzeroth, M., Megerle, A.: Renewable Energy Projects: Acceptance Risks and Their Management. Renewable and Sustainable Energy Reviews, pp. 576–584 (2012)
4. IEA.: Renewables Information. International Energy Agency, Paris (2013)
5. REN21: Renewable Energy Network for te 21th Century. Obtenido de http://www.ren21.net (2011)
6. Connolly, D., Lund, H., Mathiesen, B.: Smart Energy Europe. Renewable and Sustainable Energy Reviews, pp. 1634–1653 (2016)
7. Larsen, H., Petersen, L. S.: The future energy system distributed production and use. Denmark: Riso National Laboratory (2005)
8. DOF.: Díario Oficial de la Federación. Obtenido de http://dof.gob.mx (2006)
9. DOF.: Diario Oficial de la Federación. Obtenido de http://dof.gob.mx (2008)
10. Black, M. J., Sadhukhan, J., Kenneth, D., Drage, G., Murphy, R.J.: Developing database criteria for the assessment of biomass supply chains. Chemical Engineering Research and Design, pp. 253–262 (2016)
11. Engelken, M. et al: Comparing drivers, barriers, and opportunities of business models for renewable energies: A review. Renewable and Sustainable Energy Reviews, pp. 795–809 (2016)
12. Gatzert, N., Kosub, T.: Risk management of renewable energy projects: The case of onshore and off shore wind parks. Renewable and Sustainable Energy Reviews, pp. 982–988 (2016)
13. Yudi, F., Sofri, Y.: Challenges in implementing renewable energy supply chain in service economy era. Industrial Engineering and Service Science 2015, IESS 2015 (2015)
14. Goh, H., et al.: Renewable energy project: Project management, challenges and risk. RenewableandSustainableEnergyReviews, pp. 917–932 (2014)
15. D. Lund, P.: How fast can businesses in the new energy sector grow? An analysis of critical factors. Renewable Energy, pp. 33–40 (2013)
16. Lokey, E.: Barriers to clean development mechanism renewable energy projects in Mexico. Renewable Energy, pp. 504–508 (2008)

17. Aslani, A., Mohaghar, A.: Business structure in renewable energy industry: Key areas. Renewable and Sustainable Energy Reviews, pp. 569–575 (2013)
18. Loong Lam, H., et al.: Optimisation of regional energy supply chains utilising renewables: P-graph approach. Computers and Chemical Engineering, pp. 782–792 (2009)
19. López Torres, V. G., Alcalá Álvarez, C., Moreno Moreno, L. R.: La cadena de suministro de la energía solar. Conciencia Técnologica, pp. 18–23 (2010)
20. Mikatia, M., et al,: Modelado y Simulación de un Sistema Conjunto de Energía Solar y Eólica para Analizar su Dependencia de la Red Eléctrica. Revista Iberoamericana de Automática e Informática industrial, 267–281 (2012)
21. Rylatt, M., Gammon, R., Boait, P., Varga, L., Allen, P., Savill, M. Cascade: An agent based framework for modeling the dynamics of smart electricity systems. E:COIssue. (2013)
22. Yildiz, Ö.: Financing renewable energy infrastructures via financial citizen participation - The case of Germany. Renewable Energy, pp- 677–685 (2014)
23. Nielsem, J. H., Ehimen, E. A.: Biomass supply chains for bioenergy and biorefining. Reino Unido: Elsevier (2016)
24. Waters, D.: Logistics An Introduction to supply chain management. New York, EUA: Palgrave Macmillan (2003)
25. Rendon, G; et al.: Dynamic analysis of feasibility in ethanol supply chain for biofuel production in Mexico. Applied Energy, ELSEVIER (2014)
26. Ramos, R., et al.: Assessing the impact of a vinasse pilot plant Scale-up on the key processes of the ethanol supply chain. Mathematical Problems in Engineering. Hindawii (2016)
27. d'Amore, F., Bezzo, F.: Strategic optimisation of biomass-based energy supply chains for sustainable mobility. Computers and Chemical Engineering, pp. 68–81 (2016)
28. Azadeh, A., Vafa Aran, H.: Biodiesel supply chain optimization via a hybrid system dynamics mathematical programming approach. Renewable Energy, pp. 383–403 (2016)
29. Barlas, Y.: System dynamics: Systemic feedback modeling por policy analysis. UNESCO-EOLSS, Department-Boqvazici University , Istanbul, Turqía (2010)
30. Paiano, A., Lagioia, G.: Energy potential from residual biomass towards meeting the EU renewable energy and climate targets. The Italian case. Energy policy, pp. 161–173 (2016)
31. Ballou, R.: Logística: Administración de la CS. México, México: Pearson (2004)
32. De Laportea, A.D., Weersinka, A.J., McKenney, D.W.: Effects of supply chain structure and biomass prices on bioenergy feedstock supply. Applied Energy, pp. 1053–1064 (2016)
33. Cancino-Solórzano, Y., Villicaña-Ortiz, E., Gutiérrez-Trashorras, A. J., Xiberta-Bernat, J.: Electricity sector in Mexico. Renewable and Sustainable Energy Reviews (2009)

ISSN 1870-4069

Agile Dimensional Model for a Data Warehouse Implementation in a Software Developer Company

Kathya E. Mercado[1], Cynthia B. Perez[1], Laura P. Lopez-Arredondo[1], Karina Caro[2], Luis A. Castro[1] and Luis-Felipe Rodriguez[1]

[1] Sonora Institute of Technology (ITSON), Ciudad Obregon, Sonora, Mexico

[2] Drexel University, USA

{kathya.mercado, laurapaolalpz}@hotmail.com, {cynthia.perez, luis.rodriguez}@itson.edu.mx, luis.castro@acm.org, karinacaro@drexel.edu

Abstract. Nowadays, the increasing development of Business Intelligence (BI) solutions in organizations, has enabled executives achieve a better understanding of business information for timely and rapid decision-making in a tremendously dynamic market. Although there is an increasing interest in adopting an agile approach to the software development, the emergent need of using agile methodologies in BI solutions is undeniable. This paper discusses the importance of using agile methods in the design and development of data warehouses taking into account the business processes, requirements analysis, and organizational objectives. Thus, we present a case study derived from a real-world business project where the agile methodology Business Event Analysis and Modeling (BEAM) is used to design the data warehouse. The project is based on a billing system with about one million operations on a daily basis with more than 15,000 clients. Finally, the results of this paper include the design of the dimensional model using an agile approach, the construction of the data warehouse through the ETL processes and an interactive dashboard according to the key performance indicators defined by the business decision-makers.

Keywords: business intelligence, agile dimensional model, data warehouse.

1 Introduction

In the business world we aim to obtain greater profits and a greater competitive advantage, *hence*, appropriate timely decision-making plays a very important role in the fulfillment of the organizational objectives. Thus, the use of Business Intelligence (BI) systems can help meet these needs as it provides a set of methodologies, applications, and practices focused on the information management for accurate, timely decisions making in an organization.

A major component of any BI model is the design and implementation of a data warehouse in organizations, since it provides valuable and strategic information to support decision-making through real-time access to business transactions and advanced business analytics [1]. According to Inmon [2] and Imhoff [3], a data warehouse in organizations is an integrated data collection, non-volatile and variable over time. Hence, it has a complete history of the organization, beyond the transactional and operational information favoring the data analysis for decision-making.

Today, the two most widely used methodologies for the design and implementation of data warehouses are the model of Inmon [2] and Kimball [4]. They consider the data warehouse as the central repository of data for organizations that is used to present business reports. The difference between these two methodologies lies in how to make deliveries of progress (time) and how to manage changes during the process (see Table 1).

Recent studies tend to show that use of agile methodologies for the design and implementation of data warehouse in organizations is playing an important role to obtain value information to help decision-makers and to generate competitive advantage by improving the extraction and processing knowledge. These studies present a structured methodology, inspired from the agile development models as Scrum, XP and AP [5]. Thus, an increasing number of companies are choosing for an agile philosophy in software development due to the constant need to be flexible and adaptable to the technological changes and the new user demands [6-10].

For this reason, it is important to consider agile methods for the design and implementation of a data warehouse in a BI system (see Table 1). In this regard, it has been shown that agile development processes increase the potential for developing the success of a data warehouse by solving many of the typical problems presented in traditional methodologies [11, 8, 12].

In this paper, we use the BEAM methodology in the design, modeling, and implementation of a data warehouse for a management system. This system processes one million operations on a daily basis from more than 15,000 clients. The agile design and modeling of the data warehouse are presented through a case study of a Software Development Company located in Mexico, which offers software solutions focused on meeting the needs of development, implementation and support to any business sector.

2 Agile Dimensional Modeling

Traditional data warehousing projects follow the waterfall structure to perform dimensional modeling [13]. However, its use is increasingly unlikely and alternatives of analyzing and designing similar to those used in software development projects such as agile methods are looked for [7, 12, 14]. In this respect, the agile dimensional modeling is being considered as a solution for BI systems since it allows developers to reduce the risks that the waterfall structure could produce [11, 14]. All this is possible by adopting a highly interactive, incremental and collaborative approach to the whole analysis, design and development activities of a data warehouse, such as the agile BEAM methodology [14] (see Figure 1).

Table 1. Comparison of the Inmon, Kimball, and Corr methodologies.

	Inmonn [2] (Traditonal)	**Kimball [4] (Traditonal)**	**Corr [14] (Agile)**
Business Decisions	Strategic	Tactical	Collaborative
Scope	Product owner	Project manager	Team
Approach	Top – down	Bottom – up	Bottom – up
Objective	Deliver a robust technical solution based on proven methods	Provide a solution that facilitates the end users to consult the data	Responding to change and user needs
Data requirements	Enterprise-wide	Business process	Individual business requirement (KPIs)
Data modeling	Normalized form (3NF)	Dimension model (Star or Snowflake)	Dimension model (Star or Snowflake)
Orientation	Enterprise-wide	Business process	People
Communication	Formal	Formal/Informal	Informal
Time	Longer start-up time	Shorter start-up time	Minimal start-up time
Project schedule risk	High	High	Low
Ability to respond to change	Low	Medium	High
End users involvement	Minimal	Oscillate depending on the project	High
Cost to build	High initial cost	Low initial cost	Minimal initial cost

2.1 BEAM Methodology

Corr [14] proposes the BEAM (Business Event Analysis and Modeling) methodology, an agile data modeling method for the design and development of data warehouses and data marts. This method combines analysis and modeling techniques to meet data requirements related to business events and data modeling for database design that is easy to understand by stakeholders and also, easy to translate into logical/physical models for IT developers. The BEAM methodology involves stakeholders who think beyond their current reporting requirements by describing data stories, that is, narratives that define the dimensional details of business activities necessary to be measured. In order to obtain these data stories, data modelers ask questions to stakeholders using a framework based on the 7Ws (who, what, where, when, how many, why and how) [14]. The way to find these answers of the 7Ws and make sure they inform data warehouse design is to ask end-users about the events that are happening in their business. Therefore, the enhanced Start Schema is used to generate and show schema of physical data bases, where are they involved Data Modelers, DBAs, DBMS, ETL Developers, BI Developers and Testers. This framework is one of the main activities of the BEAM methodology because it allows discovering and modeling data requirements and thus, to construct the table of dimensions and facts of the data warehouse depicted through the star model.

According to Corr [14], the BEAM methodology has several diagrams for the analysis and design of the data warehouse model, such as: BEAM Table, Hierarchy Chart, Timeline, Event Matrix and Enhanced Star Schema.

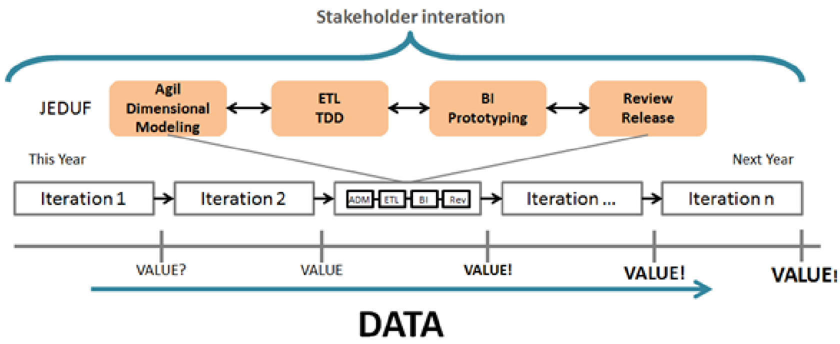


Fig. 1. Agile data warehouse development timeline. Source [14].

For instance, the BEAM Table is used to model business events and dimensions at the same time; people involved in this diagram are data modelers, business analysts, stakeholders and BI users. The Enhanced Star Schema is used to visualize the dimensional model for the implementation of the physical database schemas; the people involved in this diagram are Data Modelers, DBAs, DBMS, ETL Developers, BI Developers and Testers.

3 Agile Data Warehouse: A Case Study of a Billing System

This case study focuses on the design and implementation of a data warehouse using the BEAM methodology for a billing system of a company based in Mexico with operations in software development with around 15,000 active clients in 19 countries and more than a million operations on a daily basis. Despite all the information daily stored on the company's servers, this information is not used or analyzed so far by the working team, identifying an area of opportunity for the design and implementation of BI systems. Therefore, decision-makers could offer their customers significant knowledge through scorecards and thereby provide a competitive advantage. The management system access to a database based on the Entity-Relationship Model (ER) allowing to record, update, delete, and query information from the main business processes. This system has the following modules such as: billing, inventories, clients, payroll, branches, among others. Hence, in this case study, we focus on the billing module. The database used by the system contains around 70 tables using only the most relevant according to the key performance indicators (KPI).

3.1 Analysis, Design and Implementation of the Data Warehouse

Unlike the development of software applications, where the requirements of the organizations are often relatively well defined by the result of the stability of business rules over time; create a data warehouse depends on the reality of the company and its current conditions.

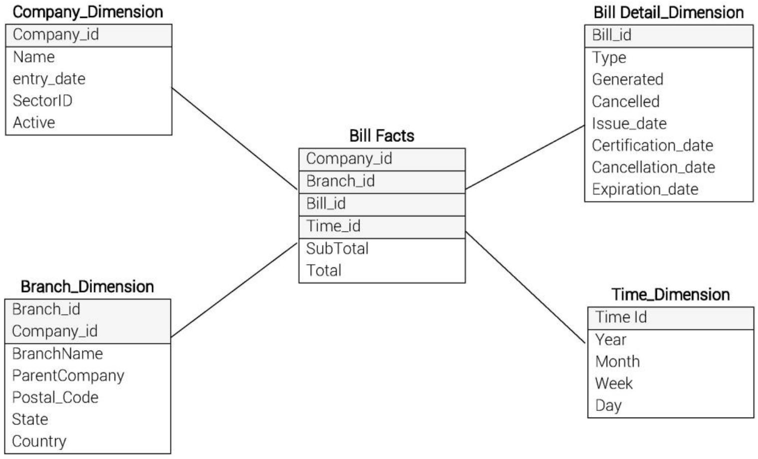


Fig. 2. Star schema for the proposed data warehouse.

Thus, the company requires identifying useful information in order to obtain significant information about its clients. In this way, the following key performance indicators (KPIs) were identified along with the business owners for the design and implementation of the data warehouse: a) Compare monthly growth of registered companies; b) Identify customer loyalty, through the use of the system more than 3 years; c) Identify quantity and list of branches by parent enterprise; d) Visualize the States where there are more than four companies using the management system; e) Measure different types of invoices volumes identifying the invoices variations not only by a time period but also, for company; f) Geographically view the top 10 companies that make the most invoices for a period of time; g) Compare the invoices by branch, time and state.

In order to model the data requirements, the 7Ws framework was used resulting in the identification of facts and dimensions in the star model, presented in Figure 2.

In this regard, the proposed dimensional model contains a fact table where the measurements or metrics of a specific event are recorded; for example, the invoice for a purchase and foreign keys referencing dimensional data tables (Company, Branch, Bill Detail and Time) which contain descriptive information. In order to implement the proposed model (see Figure 2), it is necessary to perform the ETL (Extraction, Transformation, Load) process which enables moving data from multiple sources, transform and load them into the data warehouse to analyze and thereby giving valuable information to organizations. In this way, Microsoft SQL Server Integration Services (SSIS) of Visual Studio 2015 was used to perform the ETL process. Finally, the information contained in the data warehouse was visualized through a scorecard developed using the Microsoft PowerBI tool. Thus, the KPIs were analyzed in order to identify the visual elements corresponding to each key indicator.

Once the visual elements are selected, the information about the clients is then shown in the scorecard by using queries in which decision-makers could use it interactively.

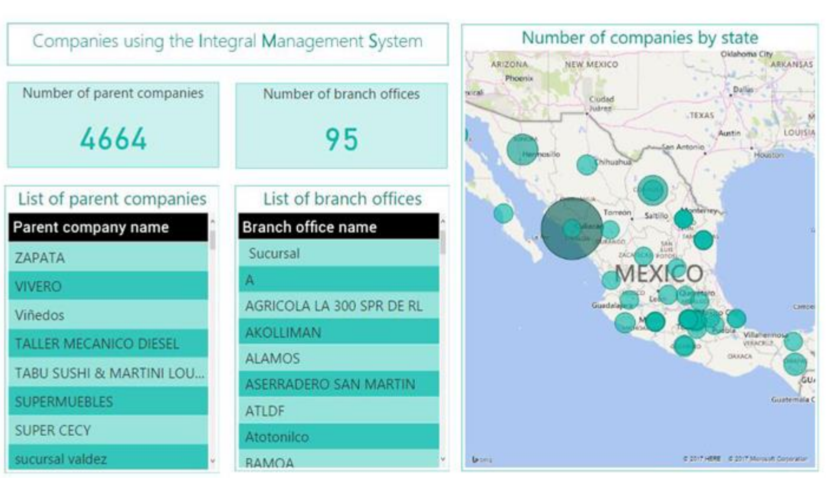


Fig. 3. Scorecard interface: Geographic coverage view of the system.

On the other hand, the software company's CEO was interested in seeing through graphs and trend lines, the situation and invoice generation behavior of those companies who use the management system.

In this way, it could be easily visualized when billing peaks are produced, i.e. sales generated by companies using the system. Figure 3 presents the view of the BI system that shows by state the number of companies who use the system, the bigger the circle the more companies that use the system are in the corresponding state. In this way, it is easy to identify the states where there is little or no presence of the system helping the decision-maker to pay attention in sale strategies. Likewise, Figure 4 presents the view related to invoice analysis where the number of invoices per type is visualized through an interactive list, a trend line to observe the behavior of the number of invoices is generated by each company who use the system and a pie chart showing the top 10 companies with the highest number of invoices issued by the system is shown.

In this way, it is possible to graphically observe the behavior of the number of invoices in a period of time by interacting with the system, allowing the decision-makers not only to know in what years more invoices per company are registered but also, detect those decreases that may indicate a risk in the strategic plan in order to develop an action plan.

3.2 Discussion of Results

The design and implementation of the proposed BI system using BEAM methodology allowed the analysis and design of the data warehouse through an agile method that focused on the users' needs and that easily respond to changes. Therefore, meetings with the working team, i.e., stakeholders, BI users, ETL developers, business analysts, among others, were held in order to compile information requirements during all stages of the project emphasizing uninterrupted communication and collaborative work.

Fig. 4. Number of invoices by time period.

On this basis, it ensures a greater understanding of the data warehouse information and the functionalities of the BI system. Moreover, the agile method maintains a logical data structure, scalable and adaptable to future functionalities such as the integration of other system's modules, predictive analysis, among others.

In this way, a robust and scalable BI system was designed and implemented where decision-makers can count on reliable, fast, flexible and easy-to-understand analyses through the scorecard, thereby facilitating the diagnosis of indicators and decision making. Accordingly, Figure 4 shows a scorecard of the invoices section, thus, the use of the scorecard provides reports of different participants in the decision-making process, representing an opportunity for homogenize and refine business processes. For this reason, it is expected to improve business opportunities through the use of key performance indicators by the extraction, processing and presentation of significant information according to the business strategic objectives. Eventually, the use of the BI system will positively impact the improvement of the company's value chain processes, its competitiveness and thus, the profitability of the business.

4 Conclusions and Future Work

Today, entrepreneurs need to analyze and interact with real-time visual information in order to support decision-making. In this regard, the methodologies used in the design of BI systems should consider the current needs and challenges where business requirements are not static and change constantly. Hence, this paper proposed the use of an agile dimensional model for the design and implementation of a data warehouse based on the BEAM methodology applied to a case study for a Software Development Company. In order to complete the project successfully, the organizational requirements were defined, the star schema was modeled, the ETL process was executed, the data warehouse was implemented and finally, the KPIs were graphically displayed into the scorecard for decision-making.

The results obtained from the use of an agile methodology are found as a model easy- to-understand for the stakeholders; for this reason, it is mandatory to involve them in the whole process.

By adopting this agile approach, flexibility is ensure, as well as, personal coordination with the stakeholders, consistency and simplicity in the whole process.

As a future work, we plan to incorporate a predictive analysis section into the BI system allowing decision-makers to discover patterns, opportunities and prevent risks by increasing the profitability of the business.

References

1. Chaudhuri, S., Dayal, U.: An overview of data warehousing and OLAP technology. ACM Sigmod record, 26(1), pp. 65–74 (1997)
2. Inmon, W. H.: Building the Data Warehousing. John Wiley & Sons (2002)
3. Imhoff, C., Galemmo, N., Geiger, J. G.: Mastering data warehouse design: relational and dimensional techniques. John Wiley & Sons (2003)
4. Kimball, R., Ross, M.: The data warehouse toolkit: the complete guide to dimensional modeling. John Wiley & Sons (2011)
5. Wilson, N. B., Edgar, C. S.: Agile Methodology for Modeling and Design of Data Warehouses-AM4DW. World Academy of Science, Engineering and Technology, International Journal of Computer, Electrical, Automation, Control and Information Engineering, 9(9), pp. 2132–2137 (2016)
6. Golfarelli, M.: From User Requirements to Conceptual Design in Data Warehouse Design. IGI Global (2010)
7. Golfarelli, M., Rizzi, S., Turricchia, E.: Modern software engineering methodologies meet data warehouse design: 4WD. In: Data Warehousing and Knowledge Discovery, pp. 66–79 (2011)
8. Collier, K.: Agile analytics: A value-driven approach to business intelligence and data warehousing. Addison-Wesley (2012)
9. Chow, T., Cao, D.: A survey study of critical success factors in agile software projects. Journal of systems and software, 81(6), pp. 961–971 (2008)
10. Dyba, T., Dingsoyr, T.: Empirical studies of agile software development: A systematic review. Information and software technology, 50(9), pp. 833–859 (2008)
11. Boutkhoum, O., Hanine, M., Tikniouine, A., Agouti, T.: Integration approach of multicriteria analysis to OLAP systems: Multidimensional model. In: Computer Systems and Applications (AICCSA), ACS International Conference on IEEE, pp. 1–4 (2013)
12. Deshpande, K., Desai, B.: Model for Assessment of Agile Methodology for Implementing Data Warehouse Projects. International Journal of Applied Information Systems, 9(5), pp. 42–49 (2015)
13. Moody, D., Kortink, M.: From enterprise models to dimensional models: a methodology for data warehouse and data mart design. DMDW, pp. 5 (2000)
14. Corr, L., Stagnitto, J.: Agile Data Warehouse Design: Colaborative Dimensional Modeling, from Whiteboard to Star Schema. UK: Decision One Press (2011)

ISSN 1870-4069

Multi-agent Model for Urban Goods Distribution

Martín Darío Arango Serna[1], Cristian Giovanny Gómez Marín[1], Conrado Augusto Serna Urán[2], Julián Andrés Zapata Córtes[3]

[1] Universidad Nacional de Colombia, Sede Medellín, Medellín, Colombia

[2] Universidad de San Buenaventura, Seccional Medellín, Medellín, Colombia

[3] Institución Universitaria CEIPA, Sabaneta, Antioquia, Colombia

mdarango@unal.edu.co

Abstract. This paper proposes a multi-agent model to evaluate the last mile distribution process in urban areas, in a network that includes a single depot, multiple customers and multiple homogeneous and capacitated vehicles. The model is based on the behaviors of the different stakeholders and their interactions, including coordination and collaboration of agents in the allocations and routing processes. The performance of the Solomon Algorithm is the starting point of the proposed model and the comparison of both results shows the advantage of using multi-agent modelling to reduce distances and number of vehicles needed for urban freight transportation.

Keywords: multi-agent model, memetic algorithm, urban freight transport, collaboration.

1 Introduction

City logistics refers to the optimization of all the logistics and transport activities in urban areas, considering the environment, traffic, security impacts and energy savings [1-3]. It implies a complex system that involves transport processes, urban dynamics, and infrastructure planning and logistics strategies for all stakeholders that should work with an integrated and coordinated methodology to achieve better results [4]. According to Wolpert and Reuter [5], the main stakeholders in Urban Goods Distribution (UGD) are: carriers, public authorities, receivers, residents and shippers, each one with different behaviors and objectives. Other stakeholders that can be included are the urban consolidation centers and road operators. As all stakeholders have very different objectives, they should work coordinated with the aim to improve the urban freight transport performance.

Multi-Agent System (MAS) Modeling is an approach used in UGD that considers different stakeholders and their behaviors and interactions in several scenarios and policies changes as a support system for strategic, tactic and operative decisions. MAS can represent the reality of the UDG, the relationships between stakeholders and their behaviors as autonomous agents with the ability to collaborate [6]. MAS can represent city logistics systems in a flexible and natural way with the aim of understand and forecast policy measures and help to increase the efficiency and reduce the externalities of city logistics [7].

In this paper, a multi-agent model that considers the stakeholders involved in a single depot network with *n* customers and *k* homogeneous unlimited and capacitated vehicles is used to evaluate the last mile distribution process in urban areas. The model includes logistics behaviors, interactions and uses a heuristic algorithm to improve the performance of UGD system.

An initial allocation and routing to last mile distribution is obtained by the Solomon heuristic that is a benchmark solution to the capacitated vehicle routing problem with time windows. These initial routes are analyzed and compared with those generated by the proposed MAS model.

2 Multi-Agent Systems in Goods Distribution

There are many strategies used in urban goods distribution [8-10]. These strategies consider different distribution structures that involve the goods flow from origin to destination and the space-temporal patterns in the vehicle routing in distribution systems. MAS modeling is a used as an approach in UGD because of its ability to consider different stakeholders, their behaviors and interactions in several scenarios and policies changes and how these changes affect the decision making process. Other important MAS characteristic is that it can handle complex systems with large numbers of heterogeneous and autonomous agents which can interact and collaborate between them [7, 12].

According to Wooldridge [13] Agents are computer systems with two important capabilities: First, they have the capacity to carry out autonomous actions; second, they can interact with other agents, not only through data interchange, but also real social relationships such as coordination, cooperation and negotiation, among others. An agent can have a collection of several actions that can be performed in the interaction with its environment, which may also include other agents. The issue to the agent is deciding what actions to take to meet their internal goals [14].

A MAS organizes agents according to their characteristics and abilities to access jointly, but also in a decentralized manner to the environment in which exist and solve common and individual objectives. For this, a series of communication, coordination and negotiation protocols is established among the agents. In this case, collaboration between the system components facilitates problems solving and reaching out the defined objectives [11-15].

The application of MAS in transport system has been used in several research studies for transportation problems, but in logistics and urban freight transport there is too much work to do [16]. Some researches that use MAS for transport are [17- 21] and for Urban Goods Distribution are [14, 22, 23-27].

3 Methodology

3.1 Urban Distribution Problem

The UGD could be performed with different strategies that implies specific structures and typologies [9]. The focus of this paper is the last mile distribution and operative decisions, where a single product is delivered to a set of customers from a unique depot or satellite. The depot has no limitation for the number of homogenous capacitated vehicles. This problem can be described as a capacitated vehicle routing problem with time windows (CVRPTW) and the objective is the minimization of total delivery cost. The mathematical model for the CVRPTW expressed by Zulvia et al. [28] is:

$$\sum_{k=1}^{h}\sum_{i=0}^{n}\sum_{j=0}^{n} c_{ij x_{ij,}^{k}} \tag{1}$$

where c_{ij} is the travel cost from customer i to j, the depot is the customer 0, and x_{ij}^{k} represent decision variable defined in (2):

$$x_{ij}^{k} = \begin{cases} 1\ if\ the\ vehicle\ k\ travel\ from\ customer\ i\ to\ j\ , \\ 0\ otherwise \end{cases} \tag{2}$$

constrained to:

$$\sum_{i=0}^{n} x_{i0}^{k} - \sum_{j=0}^{n} x_{0j}^{k} = 0,\ \ \forall k = 1, \dots, h, \tag{3}$$

$$\sum_{i=0}^{n}\sum_{k=1}^{h} x_{ij}^{k} = 1\ ,\ \ \ \forall j = 1,2, \dots, n, \tag{4}$$

$$\sum_{j=0}^{n}\sum_{k=1}^{h} x_{ij}^{k} = 1,\ \ \ \forall i = 1,2, \dots, n, \tag{5}$$

$$\sum_{j=0}^{n} x_{0j}^{k} = 1,\ \ \ \forall k = 1,2, \dots h, \tag{6}$$

$$\sum_{i=0}^{n} x_{i0}^{k} = 1,\ \ \ \forall k = 1,2, \dots, h, \tag{7}$$

$$\sum_{j=1}^{n}\sum_{i=0}^{n} q_i x_{ij}^{k} \le C,\ \ \ \forall k = 1,2, \dots, h, \tag{8}$$

$$q_i \le r_i^k \le C, \forall i = 1,2, \dots, n; k = 1,2, \dots, h, \tag{9}$$

$$\sum_{i=0}^{n} x_{ij}^{k} - \sum_{i=0}^{n} x_{ji}^{k} = 1, \quad \forall j = 1,2, \dots, n; k = 1, \dots, h, \tag{10}$$

$$r_i^k + q_j - r_j^k \leq C\left(1 - x_{ij}^k\right), \quad \forall i = 0,1, \dots n; j = 1,2, \dots, h, \tag{11}$$

$$s_j^k = \max\left\{e_j, s_i^k + u_i^k + t_{ij}\right\}, \quad if\ x_{ij}^k = 1, \tag{12}$$

$$s_j^k + u_j^k \leq 1. \tag{13}$$

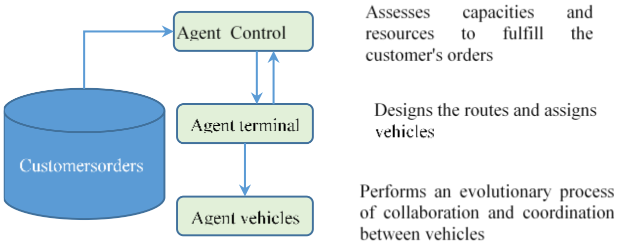


Fig. 1. Multi-agent framework to UGD.

Equation 3 is a constraint that ensures that every vehicle starts and finishes at the depot. Constraints 4 and 5 ensure that each customer is visited just once and constraints 6 and 7 ensure that each vehicle is used once. The capacity constraints are expressed in equations 8 and 9, where the demand of the customer *i* is denoted by q_i, the load of each vehicle after visiting each customer must be greater than zero r_i^k>0.

Equation 10 and 11 ensure connectivity between the created subtours. Equation 12 and 13 express the time window constraints, where e_j is the earliest time to deliver at customer *j*, l_j is the latest time to deliver at customer *j*, the beginning of the service time of the vehicle k for the customer *j* is s_j^k and u_i^k and t_{ij} are the time required to serve the customer *i* and the travel time from customer *i* to *j*, respectively.

The most common heuristic used to solve this type of problem is the Solomon insertion heuristic [29] that will be used as a benchmark to compare the proposed model solution. Although this heuristic is recognized as a good solution method, we propose a multi-agent model to improve the solutions generated by the former heuristic.

3.2 The Multi-Agent Model

The proposed framework to solve the problem of a single depot with *n* customers and *k* homogenous capacitated vehicles is a Multi-Agent Model with a memetic algorithm. It allows the collaboration and coordination between the vehicles in order to improve the routes in terms of travel distance and number of vehicles. The multi-agent model considers the demands, the time windows constraints of customers and the capacity of the vehicles for the assignment of each customer to a different vehicle.

It also considers the process of coordination and collaboration between the agents to establish routes that improve deliveries efficiency while fulfilling the customers' requirements, Fig. 1 presents the agents defined in the proposed MAS model.

The proposed MAS model integrates autonomous agents with reactive, proactive and social characteristics, based on the FIPA-ACL language (language of communication among agents) proposed by the Foundation for Intelligent Physical Agents [30] looking for improving the assignment the customers according to the available capacity in the terminal and in the vehicles, as well as design routes that make efficient deliveries in the terms of the established requirements of customers. This process involves many activities of communication and coordination between the different stakeholders of this distribution network.

The control agent has direct contact with the demands of the customer database, reads all the set of client's needs, assesses the capacities and resources of the logistics network, accepts or rejects the customers' orders and activates all the assignment process to the other agents by a Query protocol with the agent terminal. The Query Protocol allows the agent control to ask the agent terminal if it can accept the customer's orders and the agent terminal answers if it agrees with the request or rejects it.

The agent terminal receives the information and designs the initial route by the Solomon heuristic with the number of vehicles needed to perform the routes, the travel distance and the used capacity through a Request Protocol to the vehicles requesting the service of deliveries.

The agent vehicle uses an evolutionary heuristic algorithm that allows solve optimize the problem with global and local search strategies, but with the advantage of autonomy provided by the multi-agent system. Each vehicle is an agent that evolves from the interaction with other vehicles and negotiates the best possible interchange of clients to reduce the cost of the distribution [14]. Each agent has two properties: the set of clients assigned to its route, that defines its phenotype and a gender that will allow interactions with other agents and according to [31] the believes of the agents stablishes the set of genes that can crossover their genes and it's conditioned by the gender of the agents v and the fitness value *(f)* of their phenotype *l*. The believes of the agents can be denoted as the follow [14]:

$$B_i = \left\{ \begin{matrix} g_i \neq g_j \\ \left(\frac{f_i - f_j}{f_i + f_j}\right) > \alpha, \quad 0 > \propto \leq 1 \end{matrix} \right\} i, j \; are \; agents \; \in \{v_{1,} v_2, \dots v_n\}. \tag{14}$$

The first part of the equation explains that the agent only can crossover their genes with an agent of other gender. There are two types of gender: the male agents search actively a couple to crossover information; the female agents receive the corresponding bids and decides if it accepts or rejects the male agent. In the second part, α is factor of selection according to the fitness value, while the α value is closer to 0, the greater it is the probability of selecting other agent with a higher α value.

The behavior of the agent vehicles, are directly related to actions of selection, crossover and refinement of the evolutionary framework which can be explained in the following pseudo-code:

```
Algorithm 1: evolutionary heuristic
Input: male and female agents
Processing:
```

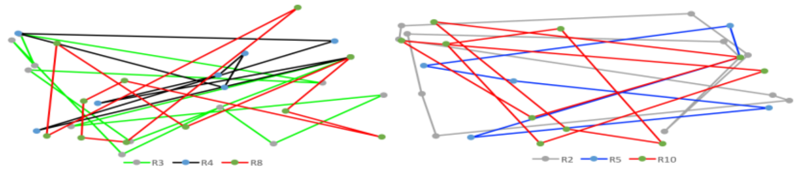


Fig. 3. Some routes of the best run. Left (Solomon) – Right (Evolutionary multi-agent).

```
(1)  Initialize.
(2)  Selection between agents according to Bi.
(3)  Crossover the genes of the agents generating N pairs
   that constitute the population of descending routes.
(4)  Assess capacity constraints of each route.
(5)  Local search with 2-opt heuristic.
(6)  Assess time window constraints.
(7)  The Children replace the Parents if their fitness
   is greater than the parents.
(8)  Randomly assigns the gender to the new individuals.
(9)  Until stop condition is fullfilled.
```

4 Model Application

For the application of the multi-agent model we consider the following sets: the demands of products are in a rank between (6 – 29) units, the location of customers is between (-100 and 100 units in a Cartesian plane, these values were chosen randomly) and the depot is located at point (69, -29); the capacity of the vehicles is 200 unit. Ten runs were performed with the same data set applying the Solomon insertion heuristic and the evolutionary multi-agent model and each run gives as a result the set of routes that should be taken to meet the demand and the time window constraints with consistent solutions. The analysis of the results for the same data demonstrates that the evolutionary multi-agent model generates an improvement of the results compared with the Solomon heuristic. The best routes produced by the Solomon insertion heuristic and by the proposed evolutionary MAS heuristic are presented in Table 1. As shown in this table, 13 routes are required to visit all the customers, and the evolutionary multi-agent model changes the sequences (must be aware that the routes start and finalize at the depot) of the routes according with the coordination and collaboration process between the agents vehicle producing an improving in the total travel distance.

The best solution for the Solomon heuristic generates a travel distance of 5632,3 units, meanwhile the best solution with the Evolutionary Multi-Agent heuristic is 5327,8 units.

This improvement in the solution, is the result of the coordination and collaboration processes of the agents' vehicle through the evolutionary algorithm. Additionally, Fig. 3 (left side) shows some of the 13 Solomon routes, allowing to observe that the routes are very heterogeneous regarding the position of customers. Due to the time window

Table 1. Best run by Solomon heuristic and Evolutionary multi-agent heuristic.

	Solomon		Evolutionary multi-agent Model	
Routes	**Sequence of customers**	**Distance**	**Sequence of customers**	**Distance**
1	85-91-65-70-81-59-76	392,273	26-60-93-66-62-20-3-9	284,871
2	82-79-99-92-90-88-83-100	284,747	34-31-44-21-80-79-70-12	390,258
3	67-56-63-61-64-62-78-75-74-53	396,85	84-65-4-82-64-15-19-48-73-47-76	316,532
4	77-58-48-22-49-45-38-41-33	531,956	22-89-88-38-51-55-6-87-56	531,955
5	89-96-93-94-87-97-84	307,937	90-83-74-13-94	326,162
6	73-54-52-50-66-72-57-60-44	485,804	11-28-49-71-32-10	260,548
7	95-98-80-86	335,186	7-39-25-98-8-58-92	380,445
8	43-32-35-19-39-46-27-25-21-18	531,799	27-96-86-68-35-52-2-5-41	526,484
9	71-51-34-26-29-16-47	547,915	53-37-99-69-78-33-23-57-46	547,807
10	40-28-31-42-37-24-11	475,728	97-29-40-43-85-61-42-18-14	477,982
11	12-4-1-3-36	395,332	67-30-45-17-24-100-36-81	479,361
12	30-7-6-8-5-15-20-2	486,721	75-63-1-95-72	419,465
13	23-14-17-10-13-9	460,136	91-59-16-77-54-50	385,989
	Total travel distance	**5.632,38**	**Total travel distance**	**5.327,86**

Table 2. Comparison between results

	Best travel distance		Mean of travel distance		Standar deviation	
Instance	Solomon	Evolutionary MAS	Solomon	Evolutionary MAS	Solomon	Evolutionary MAS
50 customers	2365,3363	2091,0110	2455,0046	2234,9492	179,5701	148,2594
100 customers	5632,3847	5327,8640	5909,3218	5371,1359	198,1098	168,1545
200 customers	5919,0469	5816,0134	6427,3153	6257,4668	221,2622	169,1766
600 customers	39302,2651	38896,9247	40941,9605	40155,5218	822,8201	732,8047

constraints, the routes are some chaotic and with superposition of trips. In the evolutionary multi-agent heuristic, the routes (right side of Fig. 3) are less complex.

To validate de results of the Evolutionary MAS, a set of instances were performed with the same heuristics. The set of instances can be found at Serna-Uran (2016) [14] for the set of 50 customers and 200 customers, and for the 600 customer the Solomon benchmark heuristic S-R1-600 was used.

Table 2, presents results of the performance of both algorithms applied in different instances and reinforces that the proposed MAS generates a lower travel distance and a lower variability in the results obtained in each run.

It is remarkable that the Solomon heuristic is a well-known benchmark to solve the CVRPTW, their solutions are feasible but with low efficiency. The presented evolutionary multi-agent heuristic is a novel model to solve this kind of problems and allows the interaction of the goods distribution stakeholders through the coordination and collaboration between them with the aim to find a global solution that is more realistic and efficient in terms of the use of resources and applicability to real problems. The Solomon heuristic as starting point of the evolutionary multi-agent systems allows, at least, to equal the initial routes, and the interaction between agents improve this routes by the behaviors that each agent perform. It should be noted that although the solutions are very near between them, the evolutionary MAS has less variability which means that the solutions are more consistent than Solomon, whereby it could be used in a more dynamic context such as the real world.

5 Conclusions

In this paper the capacitated vehicle routing problem with time windows (CVRPTW) was solved using the Solomon Heuristic and a Proposed Multi-Agent Model. The solutions found with the Solomon algorithm are feasible solutions to the Vehicle Routing Problem with time windows, however the space of solutions is not intensively explored from the insertion criteria that the heuristic uses.

This generates that the routes found are little ordered and coherent as can be observed in figure 4 (left side), which can be interpreted as inefficient routes. On the other hand, the solutions found with the evolutionary heuristics are more ordered, more efficient and equally feasible routes in terms of time window compliance that can help in the operational decisions in the UGD.

In the process followed in the implementation of the model, it is possible to highlight the ease with which problems of great computational complexity can be solved from the distributed computing paradigm. In urban conditions, where the complexity is not only computational but also stochastic or for the difficulty to articulate different actors, multi-agent models become an important solution alternative, since those present a great flexibility for the definition of different agents with capacities and conditions, allowing to explore and optimize coordination processes based on behaviors that can be adapted according to the real conditions or scenarios. In the logistic model presented, for example, vehicles are defined as agents, which facilitates the design of collaboration and optimization processes based on particular conditions that can be easily adjusted to urban environments.

The use of fuzzy logic in the relaxation of the constraints of time windows could reduce the cost of the routes with a penalization for violating the time windows, the use of this technique would be interesting for further research. In the same way, inclusion of the dynamic context of the UGD in the model according with the permanent changes presented at the demand and in the travel times in the cities is a future research filed.

References

1. Taniguchi, E., Thompson, R. G., Yamada, T., Van Duin, J. H.: City logistics: Network modelling and intelligent transport systems. Elsevier Science (2001)
2. Zapata-Cortes, J. A.: Optimización de la distribución de mercancías utilizando un modelo genético multiobjetivo de inventario colaborativo de m proveedores con n clientes. Tesis Doctoral, Universidad Nacional de Colombia (2017)
3. Arango-Serna, M. D., Zapata-Cortes, J. A., Serna-Uran, C. A.: Collaborative Multiobjective Model for Urban Goods Distribution Optimization. In: García-Alcaraz, J., Alor-Hernández, G., Maldonado-Macías, A., Sánchez-Ramírez, C. (eds) New Perspectives on Applied Industrial Tools and Techniques. Management and Industrial Engineering. Springer, Cham (2018)
4. Bozzo, R., Conca, A., Marangon, F.: Decision support system for city logistics: Literature review, and guidelines for an ex-ante model. Transportation Research Procedia (2014)
5. Wolpert, S., Reuter, C.: Status Quo of City Logistics in Scientific Literature: Systematic Review. Transportation Research Record, 2269(2269), pp. 110–116 (2012)
6. Arango-Serna, M. D., Serna-Uran, C. A., Alvarez-Uribe, K. C.: Collaborative autonomous systems in models of urban logistics. DYNA, 79(172), pp. 171–179 (2012)
7. Anand, N., van Duin, J. H. R., Tavasszy, L.: Framework for Modelling Multi-stakeholder City Logistics Domain Using the Agent based Modelling Approach. Transportation Research Procedia, 16(March), pp. 4–15 (2016)
8. Rushton, P., Croucher, P., Baker, P.: The handbook of logistics and distribution management, 3rd edition. Ed. Kogan Page Limited (2010)
9. Estrada, M. Á.: Análisis de Estrategias Eficientes en la Logística de Distribución de Paquetería. Universidad Politécnica de Cataluña (2007)
10. Arango, M. D., Zapata, J. A., Andres, C.: Metaheuristics for goods distribution. Proceedings of 2015 International Conference on Industrial Engineering and Systems Management (IESM), IEEE Publications. pp. 99–107 (2015)
11. van Lon, R. R. S., Ferrante, E., Turgut, E., Wenseleers, T., Berghe, G., Holvoet, T.: Measures of dynamism and urgency in logistics. European Journal of Operational Research, 253(3), pp. 614–624 (2016)
12. Kammoun, H. M., Kallel, I., Casillas, J., Abraham, A., Alimi, A. M.: Adapt-Traf: An adaptive multiagent road traffic management system based on hybrid ant-hierarchical fuzzy model. Transportation Research Part C, 42, pp. 147–167 (2014)
13. Wooldridge, M.: An Introduction to MultiAgent Systems. (J. Wiley, Ed.) (2001)
14. Serna-Uran, C. A.: Modelo multi-agente para problemas de recogida y entrega de mercancías con ventanas de tiempo usando un algoritmo memético con relajaciones difusas. Doctoral Thesis. Universidad Nacional de Colombia (2016)
15. Corchado, J. M.: Modelos y Arquitecturas de Agente. In Mas.A (eds). Agentes software y sistemas multi-agente: conceptos, arquitectura s y aplicaciones. Prentice Hall (2005)
16. Taniguchi, E., Thompson, R. G., Yamada, T.: Emerging Techniques for Enhancing the Practical Application of City Logistics Models. Procedia - Social and Behavioral Sciences, 39, pp. 3–18 (2012)
17. Bernardi, D., Confessore, G., Stecca, G. A.: Multi-Agent Model Integrating Inventory and Routing Processes. Virtual Enterprises and Collaborative Networks, 149, pp. 107–114 (2004)
18. Sitek, P., Wikarek, J., Grzybowska, K.: A Multi-Agent Approach to the Multi-Echelon Capacitated Vehicle Routing Problem. In Highlights of Practical Applications of

Heterogeneous Multi-Agent Systems. The PAAMS Collection pp. 121–132 (2014)

19. Kohout, R., Erol, K.: In-time agent-based vehicle routing with a stochastic improvement heuristic. Proceedings of the 6th National Conference on Arti cial Intelligence (AAAI-99), pp. 864–869 (1999)
20. Zargayouna, M., Balbo, F., Scémama, G.: A multi-agent approach for the dynamic VRPTW. International Workshop on Engineering Societies in the Agents World, pp. 1–15 (2008)
21. Vokrinek, J., Komenda, A., Pechoucek, M.: Agents Towards Vehicle Routing Problems. In Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems, pp. 773–780 (2010)
22. Arango-Serna, M. D., Serna-Uran, C. A., Zapata-Cortes, J. A.: Multi-agent System Modeling for the Coordination of Processes of Distribution of Goods Using a Memetic Algorithm. In: García-Alcaraz J., Alor-Hernández G., Maldonado-Macías A., Sánchez-Ramírez C. (eds), New Perspectives on Applied Industrial Tools and Techniques. Management and Industrial Engineering. Springer, Cham (2018)
23. De Oliveira, L. K., Lessa, D. A., Oliveira, E., Ferreira, B., Calazans, G.: Multi-agent modelling approach for evaluating the city logistics dynamic in a vulnerability situation: An exploratory study in Belo Horizonte (Brazil). Transportation Research Procedia, 25, pp. 1046–1060 (2017)
24. Schröder, S., Dabidian, P., Liedtke, G.: A conceptual proposal for an expert system to analyze smart policy options for urban CEP transports. In 2015 Smart Cities Symposium Prague, SCSP´15 (2015)
25. Cavalcante, R. A., Roorda, M. J.: Freight market interactions simulation (FREMIS): An agent- based modeling framework. Procedia Computer Science, 19, pp. 867–873 (2013)
26. Joubert, J. W.: Analyzing Commercial Through-Traffic. Procedia - Social and Behavioral Sciences, 39, pp. 184–194 (2012)
27. Baindur, D., Viegas, J. M.: An agent based model concept for assessing modal share in inter-regional freight transport markets. Journal of Transport Geography, 19(6), pp. 1093–1105 (2011)
28. Zulvia, F. E., Kuo, R. J., Hu, T. L.: Solving CVRP with time window, fuzzy travel time and demand via a hybrid ant colony optimization and genetic algortihm. 2012 IEEE Congress on Evolutionary Computation, CEC 2012, pp. 10–15 (2012)
29. Ghiani, G., Laporte, G., Musmano, R.: Introduction to Logistics Systems Planning and Control. West Sussex: Wiley (2004)
30. Foundation For Intelligent Physical Agents. Retrieved from (2015)
31. Wang, H., Zeng, J., Xu, Y., Meng, A., Ye, L., Roy, D., Vacher, J..: Genetic algorithms using multi-objectives in a multi-agent system. Robotics and Autonomous Systems. 33, pp. 179–190 (2005)

ISSN 1870-4069

Design and Implementation of a Data Warehouse to Support Decision-Making in a Health Environment

Simon G. Cornejo[1], Karina Caro[2], Luis-Felipe Rodriguez[1], Roberto Aguilar A.[1], Cynthia B. Perez[1], Luis A. Castro[1]

[1] Sonora Institute of Technology (ITSON), Ciudad Obregon, Sonora, Mexico

[2] Drexel University, Philadelphia, Pennsylvania, USA

{scornejo175319 raguilar175318}@alumno.itson.edu.mx, karinacaro@drexel.edu, {luis.rodriguez, cynthia.perez}@itson.edu.mx, @itson.edu.mx, luis.castro@acm.org

Abstract. One of the most common challenges in the management of electronic medical records (EMR) is to extract critical knowledge that could serve to enhance collaboration among medical doctors to support decision-making. In this paper, we present the design and implementation of a data warehouse designed to strengthen decision-making related to epidemiological patterns, trends, areas of influence and other pathologies. The data warehouse is based on a relational model that contains information about EMR of different states of Mexico. We used the Business Event Analysis and Modeling (BEAM) methodology for the design and implementation of the data warehouse, a novel methodology to design agile data warehouses. Using BEAM, we created a star dimensional model and defined the Extract, Transform and Load (ETL) processes to transfer the data to the new model. In order to show the potential of our data warehouse, an interactive dashboard with different indicators was built. We close discussing how the medical doctors could use our data warehouse to support the decision-making process.

Keywords: decision support systems, data warehouse, health environment, BEAM, ETL.

1 Introduction

In recent years, the use of electronic medical records (EMR) has increased considerably, changing the way traditional records are stored and managed [1]. In EMR, the data is managed and stored digitally, enabling health professionals to maintain information in one place. The use of these systems has enabled the accumulation of large amounts of data, opening up opportunities for analyzing and obtaining relevant information from them [2].

However, in most cases, besides data queries, these data are not analyzed, leaving aside all the knowledge that can be obtained from them. These large amounts of data and the knowledge that can be obtained from them open the possibility of supporting decision-making in the health environment, as well as promoting communication and collaboration among health professionals. Having a tool to support decision-making that could provide knowledge based on patients' historical data stored in EMR, might help to a great extent to improve health services.

An alternative to support this problem is the use of a data warehouse, a repository that preserves the historical context to accurately assess an organization's performance over time [3]. It is optimized for high-throughput queries, since user queries often require that hundreds or thousands of transactions are searched for and compressed into a set of responses. Data warehouses are the basis of the processing of Decision Support Systems (DSS). They facilitate the analysis since all data are concentrated in a single data source, which can integrate both structured and unstructured data, with different granularity. A data warehouse based on the data of an EMR could support health professionals to obtain relevant knowledge to enhance the decision-making process. However, the design and construction of a data warehouse is not an easy task. First, the preparation and cleaning of large amounts of data present some challenges such as the concentration of heterogeneous data, incomplete records, and integrity errors, among others. On the other hand, the design of the data warehouse has to be done in such way that the response time of the queries is fast and the results are correct.

An optimal design of the data warehouse is required to facilitate the extraction and analysis of the stored data. The use of agile methodologies (e.g., SCRUM [4]) have presented multiple advantages in software development such as customer satisfaction by the rapid and continuous delivery of useful software. Agile methodologies are characterized by emphasis on stakeholders and interactions rather than processes and tools [4]. On the side of the design of data warehouses, the Business Event Analysis and Modeling (BEAM) methodology [5] offers several advantages for designing data warehouses. For example, individuals and interactions over processes and tools, working software over comprehensive documentation, and customer collaboration over contract negotiation. BEAM upholds these values and the agile principle of data warehouse practitioners to work directly with stakeholders to produce data models rather than requirements documents, and working Business Intelligence (BI) prototypes of reports / dashboards rather than mockups.

This paper presents the design and implementation of a data warehouse using BEAM methodology, which information was obtained from EMR. This data warehouse provides information in the form of indicators that summarize what is happening in a health environment and support appropriate and timely decision making through the identification of diseases by geographic zones, diseases by stage of life and diseases by season of the year.

The illustration of the usefulness of the designed data warehouse, following, we present a usage scenario where health professionals use the information they visualize in an interactive dashboard to support decision making:

A health professional is concerned that many patients are presenting with a rare disease in their geographical area and he is not sure that the medications he is prescribing to his patients are the most appropriate. The health professional reviews the information provided by an interactive dashboard and he realizes that the disease which he is dealing with is very common in another geographic area of Mexico. Thus, the health professional starts to get in touch with other health professionals in that geographical area to ask for opinions and share experiences in the treatment and intervention of that disease.

2 Related Work

Research contributions have been made in the areas of data warehouse design, data staging for ETL processing, data quality assurance, and healthcare data warehouse applications, mainly in developed countries, such as the United States. Existing EMR [6] data are made available in a standardized and interoperable format, thus opening up a world of possibilities for semantic or concept-based reuse, consultation and communication of clinical data. The Community Health Applied Research Network (CHARN) [7], in the United States, represents more than 500,000 patients from diverse safety nets in 11 states, aims to create a national and centralized data warehouse with multiple partners from the Center for Community Health using different EMR systems.

The work [8] describes a virtual data warehouse (VDW) of the Health Maintenance Organization Research Network (HMORN), a public, research-centric data model implemented in 17 health care systems across the United States. At the Catholic Health Initiatives research institute, data consultation tools [11] are implemented to enable end users to access the VDW for simple consultation and research readiness activities, capture for collection of study-specific data and results reported by the patient. On the other hand, the decision support system [9] based on multi-criteria data analysis – Annalisa, is an online decision support tool for individuals and clinicians interested in making a shared decision.

In Mexico, there have been initiatives and programs of innovation and technological development of the public and private sectors that have begun to get involved in the subject of e-health. In the early 1990s, the State Basic Information System (SEIB) was centralized and encompassed all 32 states by The Ministry of Health (SSA, for its acronym in Spanish). In 1995, the National Epidemiological Surveillance System (SINAVE) was created. Its coordination is carried out by SSA and it is supported by the Single Information System for Epidemiological Surveillance (SUIVE).

In 2007, SSA initiated the development of the Mexican electronic clinical record standard. In 2011, a study was conducted in Mexican Institute of Social Security (IMSS, for its acronym in Spanish) [10] to develop Quality of Care Indicators (QCI) for Type 2 Diabetes Mellitus (T2DM). The goal was to determine the feasibility of constructing QCI using IMSS's EMR data and assessing the Quality of Care (QC) provided to IMSS patients with T2DM. As a result of this study [10], 18 QCIs were developed, of which 14 were possible to construct using available EMR data. ICQs comprised both the care process and health outcomes.

The related work shows that there is a potential of using EMR to enhance health care services. However, there are few studies that use the EMR data to obtain knowledge that could support the decision-making process in the healthcare environment. In this work, we propose to use EMR data to design and implement a data warehouse aimed at supporting the decision-making in the healthcare environment. In the next section, we describe the BEAM methodology, following by the results of the design of the data warehouse and its implementation.

3 BEAM Methodology

We used the Business Event Analysis and Modeling (BEAM) methodology to design the proposed data warehouse. To the best of our knowledge, BEAM is one of the most recent and the first agile methodology in the area of Data Warehouse and Business Intelligence (DW / BI) [5]. The BEAM methodology comprises a set of collaborative techniques for modeling BI data requirements and translating them into dimensional models on an agile time scale. Among the techniques used by the BEAM methodology are the 7W's Framework, BEAM*tables, Event Matrix and Enhanced Star Schema.

3.1 7W's Framework and BEAM*tables

The 7W's framework uses questions about who, what, where, when, how many, why, and how[5], data modelers design the model by asking BI stakeholders to tell data stories using these questions. BEAM uses tabular notation and data stories to define business events in a format that is easily recognizable and understandable to BI stakeholders. It uses spreadsheets that enable an easy translation into detailed star schemas.

BEAM*tables help engage BI stakeholders to define reports that answer their specific business questions. They are used to define fact and dimension tables, and they use natural language enable BI stakeholders easily imagine, sort, and filter the low-level detail columns of a business event using the top-level dimensional attributes. BEAM*tables can describe facts, events in terms of measures, and dimensions, descriptions of the facts, which can be used to filter, group and aggregate measurements.

3.2 Event Matrix

The Event matrix documents the relationships between all events and dimensions within a model. Event matrices record events in value chain sequences and promote the definition and reuse of conformed dimensions through dimensional models.

3.3 Enhanced Star Schema

A star schema consists of a central fact table surrounded by a series of dimension tables. The fact table contains facts: the numerical (quantitative) measures of a business event. Dimension tables contain mainly textual (qualitative) descriptions of the event and provide the context for the measurements. Enhanced star schemas are standard star schemas that use BEAM short codes to record dimensional properties and design techniques that are not directly supported by generic data modeling tools.

In the following section, we describe the obtained results of applying the above BEAM techniques to design a data warehouse, as well as, the data warehouse' implementation is described.

DiseasesxVectorxMunicipality [RE]

Patient	Disease	Municipality	State	Date	Amount
[who]	[what]	[where]	[where]	[when]	[how many]
Jaime Rodriguez	Peste bubónica	Cajeme	Sonora	29/01/2011	1
María Luisa Aceves	Tifus	Salina Cruz	Oaxaca	22/06/2013	3
Fernando López	Tifus	El Fuerte	Sinaloa	20/01/2017	3
José García	Tifus	Ahome	Sinaloa	28/03/2017	3
Luis Soto	Peste neumónica	Huetamo	Michoacan	12/03/2017	1

Fig. 1. Event to show diseases by municipalities and states of Mexico.

TREATMENT [TF]

PRESCRIPTION_ID	PATIENT	DISEASE	DOCTOR	MEDICAL
	[who]	[what]	[who]	[what]
74525	Paciente 49697	Meningitis viral, sin otra especficacion	Medico 21263	AUTRIN 600
74525	Paciente 49697	Meningitis viral, sin otra especficacion	Medico 21263	BENAXIMA
70091	Paciente 105454	Fiebre exantematica enteroviral [exantema de boston]	Medico 21115	BLEMIL PLUS ARAC
73050	Paciente 218889	Vertigo epidemico	Medico 21354	DIODOQUIN
74404	Paciente 105507	Otras infecciones virales del sistema nervioso central	Medico 13388	INVIRASE
317922	Paciente 105546	Otras infecciones virales del sistema nervioso central	Medico 21374	INVIRASE
70909	Paciente 105569	Otras fiebres virales transmitidas por mosquitos	Medico 20873	
361269	Paciente 218936	Fiebre del valle del rift	Medico 21181	PAMIGEN
74729	Paciente 218956	Fiebre del valle del rift	Medico 22178	PAMIGEN

Fig. 2. Fact table that contains the IDs that relate the dimensions.

	Dimensions	Importance	Estimate	Doctor	Patient	Medical	Disease	Colony	Municipalit	State	Treatment	Prescription	Doctor	Patient
EVENT (who does what)				who		what		where			why & how		Stakeholder group	
Importance														
Estimate														
Evaluate the patient	5	300		✓	✓			✓	✓	✓			✱	✓
Diagnose the patient	7	200		✓	✓		✓	✓	✓	✓	✓		✱	✓
Give medical treatment	6	100		✓		✓		✓	✓	✓		✓	✱	✓

Fig. 3. Event matrix.

4 Results

The data repository used to design and implement the data warehouse is a relational database derived from EMR system, containing 316,295 records of medical consultations from different patients living in different cities and states of Mexico. The structure of this relational database was analyzed and only the required tables to create the data warehouse with the dimensional model obtained using BEAM were extracted.

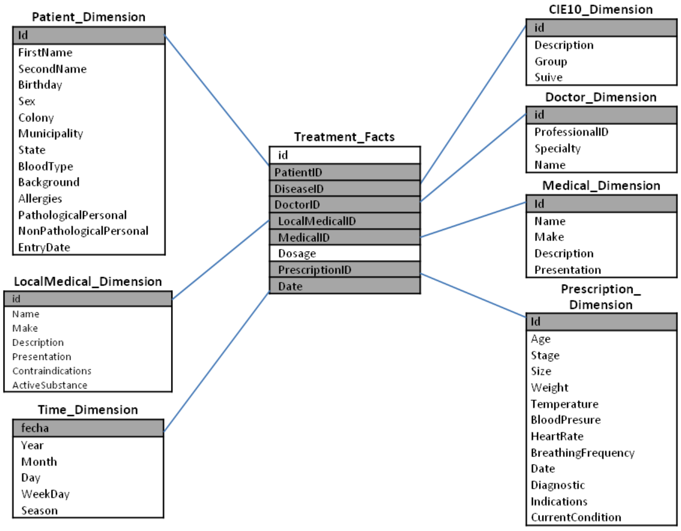


Fig. 4. Star Schema showing the dimensional model of the data warehouse.

Fig. 5. Interactive dashboard connected to the dimensional model of the data warehouse.

The tables were identified according to the indicators that the stakeholders (doctors from Obregon city identified and prioritized the indicators) defined in the business events of the BEAM methodology, which would be useful for decision making. The indicators are disease by stage of life, diseases by season of the year, epidemiology surveillance diseases, all can be filtered by municipality and state. The relational database consists of 172 tables. The tables that contain the required data to create the business events are only 7. These tables record data about patients, doctors, medications, diseases, treatments of the patients and the prescriptions that the doctors issue at the end of each medical appointment.

Once the relational database was analyzed, following the BEAM methodology, the 7W's framework technique was applied, several BEAM*tables and the event matrix were created and finally the star dimensional model was created.

7Ws Framework and BEAM*Tables. Events were defined with the help of stakeholders. In our case, the stakeholders are healthcare professionals. Examples of the defined events are Disease by stage of life, Disease by season of the year, Epidemiology surveillance diseases by municipality and state. The event shown in Fig. 1 is an example of a query to display records of patients' diseases, with granularity by municipalities and states of Mexico.

Fig. 2 shows an example of the fact table of the model that contains the IDs of the prescriptions, patients, diseases, doctors, medicines. All the descriptions corresponding to the IDs of this table are found in the derived dimension tables. This can be observed in the star dimensional model (Fig. 4), where the relationship between the IDs of the fact table and the dimension tables is defined.

Event matrix. Fig. 3 shows the event matrix with the events that happen during a medical consultation and relates them to the data of the dimensions that are involved in the development of the medical consultation.

Star dimensional model. As a result of using the BEAM techniques, we obtained the star dimensional model (Fig. 4). The dimensional model has seven dimension tables and a fact table, which contains the IDs that relate the facts to each of the dimensions. Patient dimension contains the records of the registered patients. Medical dimension contains the names, and specialty of health professionals. Medication and Medication Local dimensions are the Descriptions and brands of prescription drugs. CIE10 dimension includes the diseases and their classification SUIVE, in case of being considered epidemiological. Prescription dimension contains the symptoms of the patients at the time of the medical appointment. Time dimension is tied to the dates of the facts recorded and it enables to achieve the required granularity for the events. The relationships between the fact and the dimension tables enable agile and efficient queries, compared with the E-R model.

Validation. In order to validate the functionality of the dimensional model of the data warehouse, we performed the Extraction, Transformation and Load (ETL) processes of the data, as well as we created an interactive dashboard that shows the indicators resulting from the queries to the data warehouse.

ETL processes. The required tables were extracted from the relational database, completed and stored in temporary tables using SQL code. Next, there is an example of the code that we used to perform the extraction of the data:

*select * into DW.dbo.paciente from mmanik_completa.dbo.paciente*

*select * into DW.dbo.municipio from mmanik_completa.dbo.municipio*

*select * into DW.dbo.estado from mmanik_completa.dbo.estado*.

After the extraction process, the extracted data were transformed using the tables and fields of interest and loaded into the tables of the dimensional database. An example of the code used to perform the transformation and load processes is the following:

select m.id,cedula,u.nombre,e.nombre as especialidad,cedula_validada insert into DW.mm.medico from medico as m left join especialidad as e on especialidad_id=e.id left join usuario as u on usuario_id=u.id.

Interactive dashboard. We designed and created an interactive TABLEAU [1] dashboard once the ETL processes were finished to validate the efficiency of the data warehouse design. This dashboard is connected with the data of the dimensional model and performs queries to extract the stored data in a rapid and simple way. The dashboard shows the diseases by stage of life presented by patients (Fig 5). The dashboard also shows the diseases by seasons of the year, with the same adjustments as the indicator of diseases by life stage. Additionally, the dashboard shows the diseases that are classified in the SUIVE in a geographical map.

Thus, the health professionals can observe the map with all the diseases or select one that is of his/her interest. All indicators can be filtered by state and municipality.

5 Conclusion

With our results it is possible to support the decision making related to epidemiological patterns, trends, areas of influence and other pathologies, based on the indicators that emerged in the business events of the proposed dimensional model.

This paper shows how the BEAM methodology can be applied to design a data warehouse based on EMR system. In the future, we plan to evaluate the use of the dashboard with health professionals to investigate the potential of the data warehouse in supporting decision making in a real-case scenario.

References

1. Berner, E. S., Detmer, D. E., Simborg, D.: Will the wave finally break? A brief view of the adoption of electronic medical records in the United States. J. Am. Med. Informatics Assoc., 12(1) pp. 3–7 (2005)
2. Abidi, S.: Knowledge Management in Healthcare 63, pp. 5–18 (2001)
3. Kimball, R., Ross, M.: The data warehouse toolkit: the complete guide to dimensional modelling (2011)
4. Skarin, M., Kniberg, H.: Kanban y Scrum – obteniendo lo mejor de ambos (2010)
5. Corr, L., Stagnitto, J.: Agile Data Warehouse Design, 1 th. Leeds, UK: DecisionOne Press (2012)
6. Marco-Ruiz, L., Moner, D., Maldonado, J. A., Kolstrup, N., Bellika, J. G.: Archetype-based data warehouse environment to enable the reuse of electronic health record data, Int. J. Med. Inform., 84(9) pp. 702–714 (2015)
7. Laws, R., Gillespie, S., Puro, J., Van-Rompaey, S., Quach, T., Carroll, J., Chang Weir, R., Crawford, P., Grasso, C., Kaleba, E., McBurnie, M. A.:The Community Health Applied Research Network (CHARN) Data Warehouse: a Resource for Patient-Centered Outcomes Research and Quality Improvement in Underserved, Safety Net Populations, eGEMs (Generating Evid. Methods to Improv. patient outcomes), 2(3), pp. 10–14 (2014)
8. Ross,T. R., Ng, D., Brown, J. S., Pardee, R., Hornbrook, M. C., Hart, G., Steiner, J. F.: The HMO Research Network Virtual Data Warehouse: A Public Data Model to Support

[1] https://www.tableau.com/

Collaboration, eGEMs (Generating Evid. Methods to Improv. patient outcomes), 2(1) (2014)

9. Dowie, J., Kjer-Kaltoft, M., Salkeld, G., Cunich, M.: Towards generic online multicriteria decision support in patient-centred health care, Heal. Expect., 18(5), pp. 689–702 (2015)
10. Pérez-Cuevas, R., Doubova, S. V., Suarez-Ortega, M., Law, M., Pande, A. H., Escobedo, J., Espinosa-Larrañaga, F., Ross-Degnan, D., Wagner, A. K.: Evaluating quality of care for patients with type 2 diabetes using electronic health record information in Mexico, BMC Med. Inform. Decis. Mak., 12(1) pp. 50 (2012)
11. Bailey, D., Weeks, J., Evans, E., Lowery, J., McFarland, L.: Technologies for Managing Virtual Data Warehouse Access and Identifying Appropriate Levels of Staffing at CHI Institute for Research and Innovation. Journal of Patient-Centered Research and Reviews, 4(3), pp. 197–198 (2017)

ISSN 1870-4069

Ontology-Based Operational Risk Identification in 3PL

J.C. Osorio[1], D.F. Manotas[1], J.L. García[2]

[1] Escuela de Ingeniería Industrial, Universidad del Valle, Cali, Colombia

[2] Universidad Autónoma de Ciudad Juárez, Departamento de Ingeniería Industrial y Manufactura, México

{juan.osorio, diego.manotas}@correounivalle.edu.co, jorge.garcia@uacj.mx

Abstract. Supply chain risk management is an important activity in current supply chain management, and operational risk is one of the most important risks in supply chains. The risk identification process is one of the most important activities in supply chain risk management system. The participation of Third Party Logistics providers (3PL) in supply chains has been increasing, and it is important to consider how their presence affects risk management. On the other hand, ontologies define a common vocabulary to share information in a domain. Considering all these aspects, we propose an ontology approach to operational risk identification in supply chain that involves third party logistics providers. The ontology-based approach is oriented to improve communications about risks through the whole supply chain, achieving better results in risks management activities. The approach is validated in ground transportation activities.

Keywords: supply chain risk management, operational risk, third party logistics, ontology, ground transportation.

1 Introduction

Supply Chain Risk Management (SCRM) has become in one of the more important activities into de Supply Chain Management (SCM), and according with [2] SCRM has emerged as an important research subject in the field of SCM. SCRM has also taken on an increased importance for firms, particularly as global sourcing has increased, companies have “leaned out” their supply chains (SC), and product cycle times have become shorter [5]. The recent trend of focusing on core competence has also contributed to the popularity of logistics outsourcing and the participation in supply chain the third party logistics companies (3PL). According to this trend and its role for the success of SC, it is important to consider that the participation of 3PL providers modifies the traditional structures of supply chains, where 3PL providers become the new links and new risks arise or their probabilities or impacts change.

SCRM consist in four components [7, 14]: risk identification, risk prioritization and evaluation, risk management and risk monitoring, and some authors are agree that the success of SCRM resides mainly in the first two steps.

The identification of operational risks is the fundamental step in risk management. According to [14], qualitative techniques are used the most, and they rely on the knowledge and expertise of the experts involved. But considering this, it is necessary to have a unified language that allows sharing information throughout the chain. Sharing risk is a fundamental activity oriented to achievement an effective SCRM.

Operational risk is one of the most important risks in SC. Operational risks refer to the inherent uncertainties in day-to-day operations, such as the uncertainty in consumer demand, uncertainty in supply and uncertainty of costs.

"An ontology defines the basic terms and relations that compose the dictionary of the field of interest and the rules that combine the terms and relations so that the dictionary of terms is extended" [13].

In this paper we propose an ontology based approach to operational risk identification in supply chain that consider 3PL companies. Our proposal is oriented to improve both, risk identification activities and sharing information about risk among the whole supply chain.

2 Literature Review

2.1 Supply Chain Risk Management (SCRM)

Risk management in the supply chain refers to the concept of Supply Chain Risk Management (SCRM), which would be beneficial to the parties involved in terms of cost reduction and increase in profitability [10]. Management of risks is becoming the focal concern of the firms to survive in a competitive business environment. Thus SCRM has emerged as a natural extension of supply chain management with the objective of identifying the potential sources of risks and proposing appropriate action plans to mitigate them. But elaborating an effective SCRM program is a critical task and requires abilities in multiple areas [25].

Considering these it is clear that managing the risks present in the supply chain is an activity that requires the interest and effort of organizations looking to keep their position in the market [14].

An effective system for supply chain risk management has to identify, evaluate and quantify risks in such a way that the organization is able to generate its plans depending on the risks that have the greatest impact on their corporate objectives [19].

A risk management system has four clearly defined phases: Identification, evaluation and prioritization, management, and finally monitoring. According to some authors the first two phases are critical for the success of the system, and the definition of priorities becomes definitive when deciding on the actions required for the identified risks to be mitigated or eliminated [20].

2.2 Operational Risk Management

Although the proper definition of operational risk has often been the subject of past heated debate there is general agreement among risk professionals that the definition should, at a minimum, include breakdowns or failures relating to people, internal processes, technology or the consequences of external events [28].

Operational risks abound in every sector of the economy and in every human endeavor. Operational risks are found much sectors and in all activities [28]. Operational risks include most of what can cause an organization harm, that is foreseeable and, to a very large extent, avoidable – if not the events themselves, then at least their impact on the organization [28].

Risk is a measure of random fluctuations in performance through time. Operational risk measures the connection between those performance fluctuations and business activities [18]. Operational risks are foreseeable, and to some measure, avoidable (if not the adverse event, at least its consequences on the organization). It is clear that operational risks might be mitigated only after they have been correctly identified. A risk that has been correctly identified is no longer a risk; it becomes a management problem [20]. That is one of the most important reasons for work in risk identification.

2.3 Risk identification

The first step of risk management is to identify the sources or drivers of risks. Enterprises need to collect all possible threats systematically. The risks can be found in different aspects, either from external environment or internal operations. In logistics chains, the chance of exposure to risk is higher than other departments [27].

In order to manage and control risk effectively, management needs a clear and detailed picture of the risk and control environment in which they operate. Without this knowledge, appropriate action cannot be taken to deal with rising problems. For this purpose, risks must be identified. This includes the sources, the events and the consequences of the risks [28].

Risk identification involves a comprehensive and structured determination of potential SC risks associated with the given problem [26]. Also risk identification should include consideration of the side effects of particular consequences. A wide range of consequences should also be considered even if the source of the risk or its cause may not be evident [7].

2.4 Third Party Logistics (3PL)

The growing need of companies to focus on the core object of their business has originated a trend towards outsourcing different activities. The activities of the supply chain have evolved from a first stage with companies responsible for their own logistical processes, to companies delegating all their logistical activities to specialized agents [14].

Successful supply chains need successful logistics, and the contemporary role of the third party logistics (3PL) has moved from simple out tasking to full outsourcing. 3PLs are not just supply chain service providers; they are now supply chain strategy partners. 3PLs work simultaneously with multiple supply chain partners[9].

The role of logistics services is very critical to conduct the smooth flow of materials and information in forward and reverse supply chain.

According to this trend and its importance for the success of supply chains, it is important to consider that the participation of 3PL providers modifies the traditional structures of supply chains, where 3PL providers become the new links and new risks arise or their probabilities or impacts change [24]. And, according to the results of the third party logistics study the most frequently outsourced activities are domestic transportation (86%), warehousing (66%), international transportation (60%), freight forwarding (44%) and customs brokerage (42%) [11]. For this reason, the proposal in this paper focuses in the transporting activities in the supply chain.

2.4 Ontologies

"An ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them. Why would someone want to develop an ontology? Some of the reasons are: To share common understanding of the structure of information among people or software agents. To enable reuse of domain knowledge. To make domain assumptions explicit. To separate domain knowledge from the operational knowledge. And, to analyze domain knowledge" [18].

An ontology is a general conceptualization of a specific domain in both human and machine readable format [13]. Ontologies can be used as a backbone for the integration of expert knowledge and the formalization of project results, including advanced predictive analytics and intelligent access to third party data, through the integration of semantic technologies [12]. Ontology has different meanings according to the community in which the concept is defined. However, in a general point of view, an ontology is used to formally describe the "nature and structure of things" in terms of categories and relationships[24] .

There are some papers where using ontologies for risk management like [1, 3, 6, 8, 13, 15-17, 21-23] but any of them are applied in operational risk management in third party logistics activities.

3 Methodology

According with some authors, there is not a single way to developing ontologies, but they agree with several points that must consider in the ontology design. These points are [4, 13, 18]:

- **Determine the domain and scope of the ontology:** that is, answer basic questions like this:

  What is the domain that the ontology will cover? It is important to know which objects are interesting for the model and which ones not.

  For what we are going to use the ontology? The same domain could be modeled with different classes according to the final objective of the ontology.

  For what types of questions the information in the ontology should provide answers? To help to domain delimitation and consider the user's point of view in the modeling process.

  Who will use and maintain the ontology? It is important to know if the in charge person know about the domain or only introducing instances.

- **Consider reusing existing ontologies:** Many ontologies are already available in electronic form and can be imported into an ontology-development environment that you are using [18].
- **Enumerate important terms in the ontology:** It is useful to write a list of all terms that we will use in the ontology. What are the terms we would like to talk about? What properties do those terms have? What would we like to say about those terms? [18].
- **Define the classes and the class hierarchy:** this step is one of the most important in the process of developing ontologies [4]. There are several approaches in developing a class hierarchy: top-down, bottom-up and a combination of them [18].
- **Define the properties of classes – slots:** Once we have defined some of the classes, we must describe the internal structure of concepts. We have already selected classes from the list of terms we created in Step 3. Most of the remaining terms are likely to be properties of these classes [18].
- **Define the facets of the slots:** Slots can have different facets describing the value type, allowed values, the number of the values (cardinality), and other features of the values the slot can take[18].
- **Create instances:** The last step is creating individual instances of classes in the hierarchy. Defining an individual instance of a class requires (1) choosing a class, (2) creating an individual instance of that class, and (3) filling in the slot values [18].

4 Results

Following the proposed methodology, these are the results for our ontology:

- **Determine the domain and scope of the ontology.**

  What is the domain that the ontology will cover? Operational risk management in supply chain with third party suppliers in the chain.

  For what we are going to use the ontology? For risk management, especially for risk identification and assessment.

  For what types of questions the information in the ontology should provide answers? Types of risks, likelihood and impact, managerial strategies for risks types, echelons in the chain.

  Who will use and maintain the ontology? Companies into the supply chain (all echelons in the supply chain with third party logistics providers)

- **Consider reusing existing ontologies:** according to the literature review, does not exists ontologies in the operational risk domain that can answer the questions presented above.

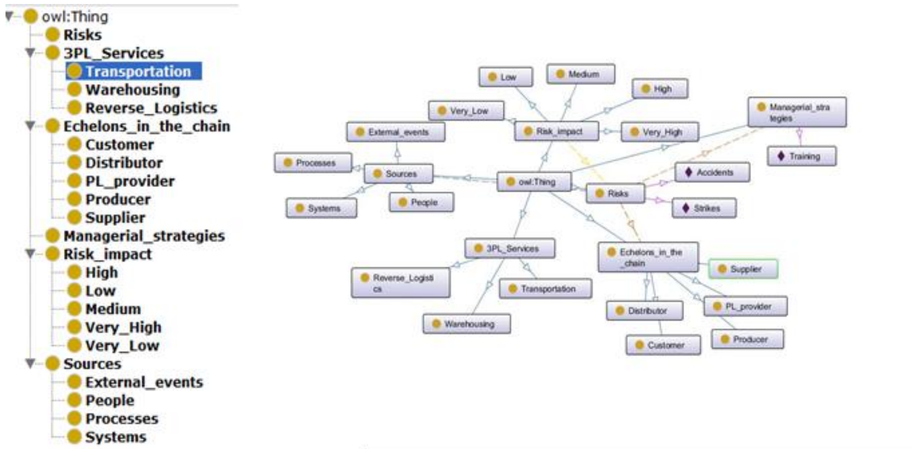


Fig.1. Ontology Classes and Class Hierarchy.

- **Enumerate important terms in the ontology:** Echelons in the chain, 3PL services, Risks, Resources, Risk impact in the strategical objectives, Management strategies.
- **Define the classes and the class hierarchy:** Figure 1 shows the classes and the class hierarchy proposed.
- **Define the properties of classes – slots:** ***belongsTo3PLServices*** (relationship between Risk and 3PL Services); ***hasRisk*** (relationship between 3PL Services and Risks); ***hasRiskImpact*** (relationship between Risks and Risk Impact); ***IsCausedFor*** (relationship between Risks and sources); ***hasManagerialStrategies*** (relationship between Risks and Managerial strategies); ***ItOccursBetween*** (relationship between Risks and Echelons in the chain).
- **Define the facets of the slots:** in Figure 2 we show the object properties of the proposed ontology.
- **Create instances:** we created some instances; particularly refer to some risks in transporting activities, such as: accidents, theft, shipping errors, strikes and driver´s lack of skills. These risks have different impact and source as show in table 1.

We hope to use the ontology in a web application that allows all echelons in the chain to share the information to improve the processes: risk identification, risk evaluation and risk management.

The idea is to expand the ontology for all 3PL activities such as warehousing, reverse logistics, picking, packing etc. And that consider different transportation modes such as maritime transportation and air transportation.

5 Conclusions

We have presented an ontology developed to improve the operational risk management in supply chain with third party logistics providers, considering that risk identification

```
############################
#   Object Properties
############################

# Object Property: :IsCausedFor (:IsCausedFor)

ObjectPropertyDomain(:IsCausedFor :Risks)
ObjectPropertyRange(:IsCausedFor :Sources)

# Object Property: :ItOccursBetween (:ItOccursBetween)

ObjectPropertyDomain(:ItOccursBetween :Risks)
ObjectPropertyRange(:ItOccursBetween :Echelons_in_the_chain)

# Object Property: :belongsto3PLservices (:belongsto3PLservices)

ObjectPropertyDomain(:belongsto3PLservices :PL_Services)
ObjectPropertyRange(:belongsto3PLservices :Risks)

# Object Property: :hasManagerialStrategies (:hasManagerialStrategies)

ObjectPropertyDomain(:hasManagerialStrategies :Risks)
ObjectPropertyRange(:hasManagerialStrategies :Managerial_strategies)

# Object Property: :hasRisk (:hasRisk)

ObjectPropertyDomain(:hasRisk :Risks)
ObjectPropertyRange(:hasRisk :PL_Services)

# Object Property: :hasRiskImpact (:hasRiskImpact)

ObjectPropertyDomain(:hasRiskImpact :Risk_impact)
ObjectPropertyRange(:hasRiskImpact :Risks)
```

Fig. 2. Object properties (Operational risk ontology).

Table 1. Instances summary.

Risks	Belongsto 3PL Services	hasRisk impact	IsCausedFor	It_Occurs Between	Managerial strategies
Theft	Transportation	High	External events	Distributor	Insurance polices
Shipping errors		Medium	Processes		New process
Strikes		Low	External events		Alternative fleet
Driver´s lack of skills		Low	People	Customer	Training

is one of the most important activities in supply chain risk management and with the proposed ontology we hope to promote sharing information about risk throughout the chain.

Although this ontology is developed to operational risk identification process, we hope in the future to involve all the operational risk management system: risk identification, risk evaluation, risk management and risk monitoring in the whole chain.

There are some papers in literature that using ontologies in risk management, but we don´t find papers using ontologies for operational risk identification in 3PL services.

References

1. Aleksy, M., Seedorf, S., Cuske, C.: A Distributed Simulation Environment for Simulation Modeling in Operational Risk Management. In: International Conference on Complex, Intelligent and Software Intensive Systems, CISIS´08, pp. 126–131 (2008)
2. Aqlan, F., Lam, S. S.: A fuzzy-based integrated framework for supply chain risk assessment. Int. J. Prod. Econ. 161, pp. 54–63 (2015)
3. Atkinson, C., Cuske, C., Dickopp, T.: Concepts for an ontology-centric technology risk management architecture in the banking industry. In: 10th IEEE Int. Enterp. Distrib. Object Comput. Conf. Work. EDOCW´06, pp. 1–7 (2006)
4. Contreras, J., Comeche, J.: Tutorial Ontologías. Madrid Univ. Complut. Madrid (2007)
5. Curkovic, S., Scannell, T., Wagner, B.: Managing Supply Chain Risk. Integrating with Risk Management. CRC Press (2016)
6. Cuske, C., Dickopp, T., Seedorf, S.: JOntoRisk: An ontology-based platform for knowledge-based simulation modeling in financial risk management. (2005)
7. Elmsalmi, M., Hachicha, W.: Risks prioritization in global supply networks using MICMAC method: A real case study. In: Int. Conf. Adv. Logist. Transp. ICALT´13. pp. 394–399 (2013)
8. Elnagdy, S. A., Qiu, M., Gai, K.: Cyber Incident Classifications Using Ontology-Based Knowledge Representation for Cybersecurity Insurance in Financial Industry. IEEE Int. Conf. Cyber Secur. In: Cloud Comput. CSCloud, 2nd IEEE Int. Conf. Scalable Smart Cloud, SSC´16, pp. 301–306 (2016)
9. Kumar, P., Singh, R. K.: A fuzzy AHP and TOPSIS methodology to evaluate 3PL in a supply chain. J. Model. Manag. 7(3), pp. 287–303 (2012)
10. Lam, H. Y.: A knowledge-based logistics operations planning system for mitigating risk in warehouse order ful fi llment. Int. J. Production Economics, pp. 1–17 (2015)
11. Langley, J., Consulting, C.: Third-Party Logistics Study : The State of Logistics Outsourcing. (2017)
12. Leibold, C., Krieger, U. H., Spies, M.: Ontology-based modelling and reasoning in operational risks. In: Kenett Ron; Yossi Raanan. (ed.) Operational Risk Management, A practical approach to intelligent data analysis, pp. 327, Wiley (2011)
13. Lykourentzou, I., Papadaki, K., Kalliakmanis, A., Djaghloul, Y.: Thibaud Latour.: Ontology-based Operational Risk Management. In: IEEE 13th Conf. Commer. Enterp. Comput., pp. 153–160 (2011)
14. Manotas-Duque. D. F., Osorio-Gómez, J. C., Rivera, L.: Operational Risk Management in Third Party Logistics (3PL). In: Alor-Hernández, G. et al. (eds.) Handbook of Research on Managerial Strategies for Achieving Optimal Performance in Industrial Processes (2016)
15. Marques, R. P., Santos, H., Santos, C.: An enterprise ontology-based database for continuous monitoring application. In: IEEE Int. Conf. Bus. Informatics, IEEE CBI´13, pp. 7–12 (2013)
16. Mohammad, M. A., Kaloskampis, I., Hicks, Y. A., Setchi, R.: Ontology-based framework for risk assessment in road scenes using videos. Procedia Comput. Sci. 60(1), pp. 1532–1541 (2015)
17. Nota, G., Rossella, A., Di Gregorio, M. P.: Ontology Based Risk Management. In: Faggini, M., Vinci, C. P. (eds.) Decision theory and choices: A complexity approach, pp. 252, Springer-Verlag (2010)
18. Noy, N. F., McGuinness, D. L.: Ontology Development 101: A Guide to Creating Your First Ontology. Stanford Knowl. Syst. Lab. 25 (2001)
19. Osorio-Gómez, J. C., Manotas-Duque, D. F., Rivera, L., García-Alcaraz, J. L.: Decision Support System for Operational Risk Management in Supply Chain with 3PL Providers. In: Alor-hernández, G., Valencia-García, R. (eds.), Current Trends on Knowledge-Based Systems, pp. 302 (2017)

20. Osorio-Gómez, J. C., Manotas-Duque, D. F., Rivera-Cadavid, L., Canales-Valdiviezo, I.: Operational Risk Prioritization in Supply Chain with 3PL Using Fuzzy-QFD. In: García-alcaraz, J. L. (eds.), New Perspectives on Applied Industrial Tools and Techniques, pp. 530 (2017)
21. Peng, G. C., Nunes, J. M. B.: Surfacing ERP exploitation risks through a risk ontology. Ind. Manag. Data Syst. 109(7), pp. 926–942 (2009)
22. Polizel, F., Casare, S., Sichman, J.: OntoBacen: A modular ontology for risk management in the Brazilian financial system. CEUR Workshop Proc. 1517 (2015)
23. Robin, C., Uma, G.: Design and Development of Ontology for Risk Management in Software Project Management. In: Proc. Int. Symp, pp. 253–257 (2009)
24. Rodriguez-Elias, O. M., Velázquez-Mendoza, M. J., Rose-Gómez, C. E.: An Ontology Based System for Knowledge Profile Management: A Case Study in the Electric Sector. In: Alor-hernández, G. and Valencia-García, R. (eds.), Current Trends on Knowledge-Based Systems, pp. 302 (2017)
25. Singhal, P., Mittalet, M. L., Agarwal, G.: Supply chain risk management : review, classification and future research directions. Int. J. Bus. Sci. Appl. Manag. 6(3), pp. 15–42 (2011)
26. Tummala, R., Schoenherr, T.: Assessing and managing risks using the Supply Chain Risk Management Process (SCRMP). Supply Chain Manag. An Int. J. 16, pp. 474–483 (2011)
27. Wee, H. M., Blos, M. F., Yanget, W. H.: Risk Management in Logistics. In: Handbook on Decision Making. Springer Berlin Heidelberg, pp. 285–305 (2012)
28. Yossi, R., Kenett, R. S., Pike, R.: Operational risk management: an overview. In: Kenett, R. S., Yossi, R. (eds.) Operational Risk Management. A practical approach to intelligent data analysis, pp. 327 (2011)

ISSN 1870-4069

The Role of ICT in the Supply Chain of Ciudad Juarez Industrial Sector

J.R. Díaz-Reza, J.L. García-Alcaraz, V. Martínez-Loya, I. Canales-Valdivieso

Universidad Autonoma de Ciudad Juarez,
Department of Industrial Engineering and Manufacturing,
Mexico

{al164440, al160648}@alumnos.uacj.mx, {jorge.garcia, icanales}@uacj.mx

Abstract. Procurement and inventory represent two of the most important stages that integrate the supply chain (SC), if there is a mistake or problem in one of them, it can lead to total failure of the chain, therefore, generate losses to the companies among other issues. To avoid that kind of problems, multiple companies have decided to invest in information and communication technologies (ICT) to improve and facilitate communication, as well as simplify the management of their inventory. In that sense, the objective of this paper is to identify the importance of ICT related to the procurement and inventory of the Mexican industrial sector, specifically, in the maquiladora industry in Ciudad Juárez, Chihuahua. To gather data, a questionnaire was developed and it was administered to that sector, obtaining 306 valid questionnaires. In addition, to measure the relationship between variables, a structural equation model was developed using partial least squares by the WarpPLS V.5® software, which integrates three hypotheses. Finally, the results indicate that ICT has a direct and positive effect on the procurement process; however, its explanatory power is greater for the inventory.

Keywords: ICT, SEM, supply chain, procurement management, inventory management.

1 Introduction

A supply chain (SC) is defined as a set of organizations that create a network where services / products, information and finance can flow. Its aim is to transform raw material into finished products which are delivered to final consumers with the highest quality, in the right quantities and at the requested time [1]. An efficient management of SC includes the ability to manage the flows of products, information and economic resources [2], thus it requires a connection between the different stages of SC to coordinate and maximize its overall effectiveness.

To achieve it, companies use ICT to improve SC efficiency, since they facilitate access to information about different agents (suppliers, transporters, manufacturers, distributors, among others), also they have played an important role in creating networks of greater value in all sectors of society [1].

The appropriate use of ICT facilitates several aspects, for example, they can allow to identify the best supplier and send him an order immediately (even automatically) [1], locate raw and finished materials in real time, providing a greater agility and visibility [3], and know inventory levels at all time.[4]. Because of the above mentioned, it is increasingly recognized that ICT are altering communication patterns between companies, customers and suppliers [5].

1.1 Problem Statement and Research Objective

Currently, the relationship between the activities of inventory management, provisioning and the use of ICT in SC is still unknown, so the objective of this work is to quantitatively measure the impact of these variables on the Maquiladora industry of Ciudad Juarez, Mexico.

2 Literature Review and Hypotheses

2.1 ICT in SC

Nowadays, companies have included ICT in their areas in order to facilitate their operations, specifically in industrial companies within dynamic areas, such as in the supply chain management [6], since ICT is believed to improve human resource, financial resource, and asset productivities, vertical integration efficiency in supply chain and customer relationship management, and competitive power in the market [7].

It is demonstrated that the use of ICT improves bilateral relations between companies and offers new opportunities for vertical collaboration, even in mature industries [8], also, the use of the internet and e-mail is common in supply chain management in all industries [9]. In order to know if a company is applying ICT properly, it is advisable to investigate the following aspects: internet use in B2B (Business-to-Business), internet use for business administration, collaboration and customization via internet, shared information and inter-organizational coordination, intra-organizational information systems for SC coordination and integration, and the use of efficient consumer response (ECR).

2.2 Procurement: Supply of Raw Materials

The acquisition of raw materials is a key activity for any company, since it facilitates to increase the creation of value minimizing the cost, this is the reason why it is crucial to have the suitable suppliers to guarantee the supply, in the right quantity at in the right quality at the right place and time. It means that the supplier selection process becomes, therefore, a crucial area of decision making [10]. In this work, in order measure the raw material supply process in a company, the following points are analyzed: cooperation level with suppliers, quantity of suppliers, use of the JIT philosophy in supply process and purchases focused on the best price.

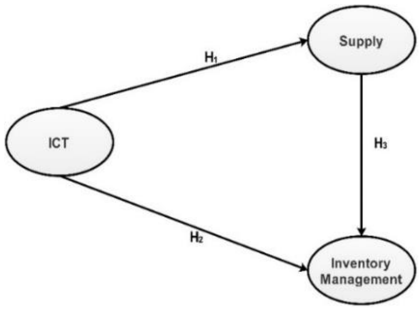


Fig. 1. Proposed model.

However, the successful adoption of ICT enables supply chain management to have a reliable information system, perform an adequate provisioning, more efficient control of inventory and material resources. Therefore, the following hypothesis is proposed

H_1: *ICT* have a direct and positive impact on the *Procurement* process in the supply chain.

2.3 Inventory Management

Inventory management is one of the factors that affect cost, which is the main reason for the focus on SC management, where coordination and collaboration between activities at all stages of the business system are considered key in the management of modern enterprises, specifically inventories consisting of raw materials, components and finished good [11]. The availability of the product at the time, in quantity, in quality and at the desired price not only provides immediate benefits, but ensures the long-term customer loyalty and brand leadership [11]. Thus, appropriate inventory management allows to develop marketing, sales, and logistics strategies, from the supplier to the final customer.

In general, in order to know whether exists an adequate inventory management in any type of company, it is necessary to identify if there is coordination and inventory management, if the automation in the warehouse is presented, if the JIT (Just in Time) is applied in delivery, if the company is focused on the low inventory costs, and if there are distribution centers nearby. In this regard, these topics are included in this research.

In that sense, ICT are a great help since they allow identifying, monitoring and transmitting information on tagged items throughout the SC, then facilitating its visibility and efficiency [2], as they increase their capacity to respond to a large quantity and variety of demanded products and meet short delivery times [2]. Therefore, the following hypothesis can be stated:

H_2: *ICT* have a direct and positive effect on the Inventory Management in the supply chain.

An important process in SC is the accurate inventory data update that adjusts and changes as goods are transferred from manufacturers to distributors [12]. However, an inaccurate inventory data update could lead to shortages and losses, so it should be avoided to maximize profit. Therefore, the following hypothesis can be constructed:

H_3: *Procurement* has a direct and positive impact on the *Inventory Management* in the supply chain.

The three hypotheses proposed for this study can be graphically represented as shown in Figure. 1, which integrates the three latent variables studied.

3 Methodology

This section describes the methodology followed to collect information and validate the three hypotheses discussed, it is integrated by distinct stages that are described below.

3.1 Stage 1: Literature Review

Multiple electronic databases, such as Springer, Emerald, Taylor & Francis, etc., are consulted and by using keywords as ICT Integration, Inventory Management and Purchasing Management essential information is identified to carry out the next stage.

3.2 Stage 2: Evaluation Instrument Design and Application

Based on information collected, a questionnaire is developed which uses a Likert scale to answer each question (1 to 5). Where 1 indicates that activity is not important or never performed, while 5 indicates that activity is always performed or extremely important [13]. The result is a questionnaire with five sections: the first includes general information of the respondent (seniority, sector, etc.). The second, seven items related to ICT; The third, four items in relation to procurement, and the last has six items in relation to the inventory.

Once the questionnaire has been developed and checked by experts, it is administered to people whose area of work is related to supply chain and it should be noted that its application is carried out by face to face interviews.

3.3 Stage 3: Analysis and Information Debugging

The information is captured in a database developed by a statistical software called SPSS 21®, where each row represents a case and the columns represent the questions. Analyzes are performed to identify missing values and if a case has more than 15% of these, it is eliminated, thus, is not considered for further analysis. If there is a smaller percentage, the missing data can be replaced by the median. Also, the extreme values are identified, so that all variables are standardized, considering as outliers to standardized values greater than the absolute value of 4. [14].

3.4 Stage 4: Questionnaire Validation

Cronbach's alpha index and the composite reliability index are used for the validation of the instrument and to measure the internal reliability by constructs [15]. Cronbach's alpha index measures the degree of correlation between items [16], accepted values are those higher than 0.7, but this value can be increased by eliminating correlated items, that probably might be explained with others [14].

Likewise, convergent and discriminant validation are performed. The first, explains the amount in which a construct converges in its indicators explaining the variance of the items [17]. On the other hand, discriminant validity indicates how different a construct is from others [17], since both concepts are very similar, the Average Variance Extracted (AVE) is used, whose acceptable values are superior to 0.5[18].

The variance inflation factor (VIF) is used to know the collinearity levels between the variables, high levels indicate that some items may be redundant[15] values below 3.3 are accepted [18]. In addition, the predictive validity is analyzed by R^2, Adjusted-R^2 and Q^2, where Q^2 is a non-parametric measure [18].

3.5 Stage 5: Structural Equation Modeling

To test the hypotheses stated in Figure. 1, the structural equation modeling (SEM) technique is used, given its ability to simultaneously examine relationships between a set of variables and including measurement error [17]. The model is evaluated in the WarpPLS software that uses partial least squares algorithms (PLS) and it is recommended for small samples without normality [18].

The quality of the model is assessed using the average path coefficient (APC), average R-squared (ARS) and average block variance inflation factor (AVIF), proposed by Kock [18]. APC and ARS are acceptable when their p-values are less or equal to 0.05, whereas for the AVIF should be less or equal to 5 [18].

The relationships between the variables are analyzed by the effects, in that context, in this model there are three types: direct (direct relations between variables), indirect (relations with two or more variables) and totals (sum of both). For each relation, the p-value is estimated in order to know the statistical significance of the effects, being the null hypothesis $\beta = 0$, against the alternative hypothesis $\beta \neq 0$.

4 Results

4.1 Descriptive Analysis

As a result, 306 valid questionnaires were collected from maquiladora industry of Ciudad Juárez. The sector that showed the highest participation was automotive (31%), followed by electric/electronic (30.7%), other (medical, aeronautical, logistics, etc.) (17%), unspecified (14.7%), plastics (4.9%) and packaging (1.6%). According to the position of respondents, it should be noted that the highest participation was from engineers (35.9%) and technicians (24.2%).

4.2 Questionnaire Validation

As described before, the Cronbach's alpha index was used, for the procurement variable was 0.776, for inventory 0.866 and for ICT 0.915, since all values were higher than 0.7 it was not necessary to eliminate any of the items.

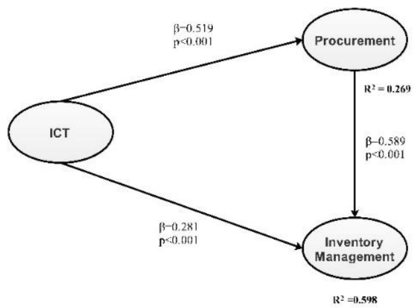


Fig. 2. Evaluated Model.

Table 1. Efficiency indexes.

Index	**Criteria**	**Value**
Average path coefficient (APC)	$p<0.05$	0.463, $P<0.001$
Average R-squared (ARS)	$p<0.05$	0.433, $P<0.001$
Average adjusted R-squared (AARS)	$p<0.05$	0.431, $P<0.001$
Average block VIF (AVIF)	Acceptable ≤ 5, Ideally ≤ 3.3	1.372
Average full collinearity VIF (AFVIF)	Acceptable ≤ 5, Ideally ≤ 3.3	2.072
Tenenhaus GoF (GoF)	Small ≥ 0.1, Medium ≥ 0.25, Large ≥ 0.36	0.520

Therefore, it is concluded that the questionnaire fulfills its objective and the information collected is valid and can be used for further analysis.

4.3 Structural Equation Modeling

Figure. 2 shows the generated model and the results obtained for the relations between the variables. The latent dependent variables show a value of R^2 allowing to establish the amount of variance explained by the independent variables. In the same way, their efficiency indices are shown in Table 1.

4.4 Direct Effects

Regarding the direct effects, the following conclusions can be established:

- **H_1**: There is enough statistical evidence to state that *Information and Communication Technologies* have a direct and positive impact on *Procurement*, since when the first variable increases its standard deviation by one unit, the standard deviation of the second latent variable also increases by 0.519 units.
- **H_2**: There is sufficient statistical evidence to point out that *Information and Communication Technologies* used along a supply chain have a direct and positive impact on the *Inventory*. When the first variable increases its standard deviation by one unit, then the standard deviation of the second latent variable increases by 0.281 units.

- **H3**: There is necessary statistical evidence to declare that *Procurement has a direct and positive effect on the Inventory,* since when the first latent variable increases its standard deviation by one unit, the standard deviation of the second latent variable also rises by 0.589 units.

Table 2. Total effects.

To	From	
	ICT	Procurement
Inventory	0.586 (p<0.001) ES=0.344	0.589 (p<0.001) ES=0.433
Procurement	0.519 (p<0.001) ES=0.269	

4.5 Indirect and Total Effects

Figure 2 shows the unique indirect effect from *ICT* to Inventory through the mediating variable *Procurement.* Its value is 0.305 and the effect size is 0.179. Finally, all the total effects are significant, since their p-values are less than 0.001 as shown in Table 2.

5 Conclusions

In this work, a model that associates three latent variables: ICT, procurement and inventory management was presented. The relationship between variables sere tested using three hypotheses, which were not rejected since their p-values are less than 0.05, because the inferences are realized at 95% confidence level.

Based on the findings, it can be highlighted that choosing the appropriate procurement methods and strategies is crucial for supply chain management because many factors that can put at risk the proper functioning of the company depends of them such as demand, delivery times, raw material, etc. so directly affecting inventories level in companies. For example, if a supply strategy is not defined and the market needs are not identified, inventory levels would decrease or increase putting in danger the commitment and capital of the company.

According to hypotheses results, it is concluded that:

1. ICT is important for have an adequate raw materials procurement and inventory management.
2. Procurement has the highest effect on inventory management and managers must pay attention to this process.

References

1. Heikkurinen, M., Appleton, O., Urciuoli, L., Hintsa, J.: Federated ICT for global supply chains: It service management in cross-border trade. IFIP/IEEE

International Symposium on Integrated Network Management (IM 2013), pp. 27–31, pp 1268–1275 (2013)
2. Ouadaa, S. E., Bah, S., Berrado, A.: Survey on supply chain ICT requirements. 10th International Conference on Intelligent Systems: Theories and Applications (SITA), pp. 20-21 pp 1–7 (2015)
3. Samdantsoodol, A., Cang, S., Yu, H., Eardley, A.: Buyantsogt, A.: Predicting the relationships between virtual enterprises and agility in supply chains. Expert Systems with Applications, 84, pp. 58–73 (2017)
4. Mensah, P., Merkuryev, Y., Manak, S.: Developing a resilient supply chain strategy by exploiting ICT. Procedia Computer Science, 77, pp. 65–71 (2015)
5. Zhang, X., Liu, Y.: The impact of ICT on supplier-buyer relationship in different types of supply chain, International Conference on Wireless Communications, Networking and Mobile Computing, pp. 21–25, pp. 4694–4697 (2007)
6. Devaraj, S., Krajewski, L., Wei, J. C.: Impact of ebusiness technologies on operational performance: The role of production information integration in the supply chain. Journal of Operations Management, 25, pp. 1199–1216 (2007)
7. Oh, S., Ryu, Y. U., Yang, H.: Supply chain capabilities and information technology characteristics: Interaction effects on firm performance. 49th Hawaii International Conference on System Sciences (HICSS), 5(8), pp. 1417–1425 (2016)
8. Albors-Garrigós, J., Hervas-Oliver, J. L., Márquez, P.: Internet and mature industries. Its role in the creation of value in the supply chain. The case of tile ceramic manufacturers and distributors in spain. International Journal of Information Management 29, pp. 476–482 (2009)
9. Donk, D. P. V.: Challenges in relating supply chain management and information and communication technology: An introduction. International Journal of Operations & Production Management 28, pp. 308–312 (2008)
10. Suryaningrat, I. B.: Raw material procurement on agroindustrial supply chain management: A case survey of fruit processing industries in Indonesia. Agriculture and Agricultural Science Procedia, 9, pp. 253–257 (2016)
11. Candan, G., Yazgan, H. R.: A novel approach for inventory problem in the pharmaceutical supply chain. DARU Journal of Pharmaceutical Sciences, 24(4) (2016)
12. Kocaoglu, M., Oksuz, C., Akan, O. B.: Effect of channel conditions on inventory database update in supply chains, 2013 First International Black Sea Conference on Communications and Networking (BlackSeaCom), pp. 252–256 (2013)
13. Likert, R.: A technique for the measurement of attitudes. publisher not identified (1932)
14. Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E.: Multivariate data analysis: Pearson new international edition. Pearson Education Limited (2013)
15. Peng, D. X., Lai, F.: Using partial least squares in operations management research: A practical guideline and summary of past research. Journal of Operations Management, 30, pp. 467–480 (2012)
16. Cronbach, L. J.: Coefficient alpha and the internal structure of tests. Psychometrika, 16, pp. 297–334 (1951)
17. Sarstedt, M., Ringle, C. M., Smith, D., Reams, R., Hair, J. F.: Partial least squares structural equation modeling (pls-sem): A useful tool for family

business researchers. Journal of Family Business Strategy, 5, pp. 105–115 (2014)

18. Kock, N.: Warppls 5.0 user manual. ScriptWarp Systems (2015)

ISSN 1870-4069

Multiobjective Model to Reduce Logistics Costs and CO_2 Emissions in Goods Distribution

Julian Andres Zapata Cortes[1], Martín Dario Arango Serna[2], Conrado Augusto Serna Uran[3]

[1] Institución Universitaria CEIPA, Sabaneta, Antioquia, Colombia

[2] Universidad Nacional de Colombia, Sede Medellín, Medellín, Colombia

[3] Universidad de San Buenaventura, Seccional Medellín, Medellín, Colombia

julian.zapata@ceipa.edu.co

Abstract. Goods transportation has increased in recent years due to new and more intensive distribution processes, such as door-to-door distribution generated by e-commerce and other marketing and logistics strategies. Transportation processes generates negative impacts in society and environment since it produces traffic jams and pollution. This paper presents a multiobjective model that simultaneously optimizes freight transportation and inventory quantity through collaboration between customers and suppliers, and also considers the distribution process CO_2 emissions. With this model decision makers in logistics can find a suitable combination between logistics costs and pollutants emission reduction. This model is solved using a multiobjective genetic model based on the NSGA II algorithm.

Keywords: goods distribution, multiobjective model, genetic algorithm, collaborative inventory, CO_2 emissions.

1 Introduction

The highly dynamic transportation processes generated by new marketing processes and changes in consumers habits have been studied for several years usually pursuing their optimization through models such as the Traveling Sales Problem (TSP) or the Vehicle Routing Problem (VRP) [1, 2]. However, these transport processes not only impact companies´ economics, but also society and the environment since it generates congestion as well as physical and chemical pollution. Therefore, professionals and academics in this area are interested in the search for processes that will improve both, economics as well as social/environmental conditions for companies and society [3].

Many authors have argued that collaboration among supply chain participants is one of the main strategies to reduce goods distribution cost, highlighting the Vendor Managed Inventory-VMI as one of the most important ways in which companies can collaborate [4]. Through VMI the inventory quantity for multiple companies can be optimized, allowing for distribution systems configurations with higher efficiency. This effectively reduces costs and transport activities intensity, as a consequence of a better inventory allocation [5]. This can be done with the Inventory Routing Problem -IRP optimization model, which, based on the collaborative inventory, allows transportation and inventory costs to be simultaneously reduced [6].

This paper analyzes the effect that inventory collaboration has on CO_2 emissions in distribution processes. Inventory and transportation in a distribution network are optimized using a multiobjective model with tree objective functions, namely: inventory cost, transport costs and CO_2 emissions. This model is based upon customers and suppliers´ collaboration under the Vendor Managed Inventory (VMI) strategy. In order to analyze the proposed multiobjective model benefits, the results are compared with the single transport optimization process through the Vehicle Routing Problem (VRP).

2 Inventory Collaboration and Optimization

Collaboration in logistics and supply chain is understood as the joining efforts of several organizations seeking superior benefits than those achievable by acting separately. For this, companies cooperate in processes such as transportation, inventory management, storage, facility design, information exchange and other logistics activities [1, 7, 8]. Since many years, supply chain collaboration has been established through approaches such as Quick Response (QR), Efficient Customer Response (ECR), Continuous product Replenishment (CPR), Vendor Managed Inventory (VMI), Planning, Collaborative Forecasting and Replenishment (CPRF) and Centralized Inventory Management, among others [1, 9, 10]. According to Díaz-Batista and Pérez-Armador [11], inventory collaboration produces a lower total annual cost than when companies work individually, generating performance improvements in the entire supply chain [12-14]. The main problem is the inventory allocation and transportation, which has been studied by multiple authors [15-17]. The most used strategy for it is the VMI [18]. The joint assignment of inventory and transportation can be done by using the IRP model [6, 19-23] as well as with multiobjective optimization approaches.

The multiobjective optimization models must be solved using complex procedures, the most widely used methods are: MOGA (Multi- Objective Genetic Algorithm), NSGA y NSGA-II (Nondominated Sorting Genetic Algorithm), SPEA y SPEA2 (Strength Pareto Evolutionary Algorithm), PAES (Pareto Archived Evolution Strategy), PESA (Pareto Envelope-based Selection Algorithm), MO-VNS (Multiobjective Variable Neighbourhood Search), DEPT (Differential Evolution with Pareto Tournaments), MO-TLBO (Multiobjective Teaching-Learning-Based Optimization), MOABC (Multiobjective Artificial Bee Colony), among others [24-28].

3 Related Works

Many authors have studied the effect of using multiobjective approaches to distributions problems. A lot of research in multiobjective transportation problems is available, however much less for models considering inventory and transport together [24]. Some works that integrated transportation and inventory management through multiobjective approaches are:

Seferlis and Pechlivanos [23] propose a model to minimize inventory level and maxmise the difference between generated revenues and associated costs. Chen and Lee [29] presented a four-objective model to optimize profits, safe inventory levels, customer service and robustness under demand uncertainties. By solving a multi-product and multi-time period production/distribution planning decisions problems, Liang [30] minimizes the total costs and total delivery time. Liao et al. [31] proposed a model for Minimizing total costs and maximizing demand satisfaction and response level. Azuma et al. [32] and Azuma et al. [33] present a model aiming to minimize transport and inventory costs using the IRP; Shankar et al. [34] propose a three-echelon capacitated plant location-distribution network in order to minimize total costs and maximize demand fulfillment.

Afshari et al., [35] minimizes the total cost of transportation, establishment/facility location, and inventory management, as well as customer satisfaction in a multi-period, multi-commodity, distribution-service network, Nekooghadirli et al. [36] minimize the costs and the average delivery time. Pasandideh et al. [37] propose a Multi-product multi-period three-echelon model that minimizes total costs and maximizes the amount of product sent to customer. Pasandideh et al. [38] similar to the later work, minimizes the mean and the total cost variance in a Supply Chain network. Zapata [28] and Arango et al [39] presented a multiobjective model to minimize inventory and transport costs, service level and required trips thorugh collaboration in a suppliers and customers network. Arango and Zapata [40] minimize transportation Costs, Inventory Costs and Service Level using the IRP. In most of the before mentioned works, authors were interested only in the companies' economics, since the goal of their research was to improve company performance or customer satisfaction.

Only Zapata [28] and Arango et al. [39] mentioned the reduction of trips required as a measure to mitigate transportation negative impact. Furthermore, these models optimize inventory and transport cost as a sum of both magnitudes, which may result in lower costs for companies but adverse environmental effects, such as an increase in trips or higher pollutants emissions. The model proposed in this article presents a new approach in which inventory and transport costs are treated apart while separately considering CO2 emissions. The aim is to evaluate different transport and inventory relationships and mitigate their environmental effect.

4 Methodology

With the aim of minimizing CO_2 emissions caused by the goods distribution process, a multiobjective model using a VMI background is proposed.

Table 1. Input parameters Customers.

Customer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Demand each Period	32	36	91	52	76	10	85	79	22	36	68	46	55	65	73
Inventory Cost	0.02	0.03	0.03	0.02	0.02	0.03	0.04	0.04	0.02	0.04	0.02	0.02	0.02	0.03	0.02
Initial Inventory	32	72	182	52	152	20	85	79	22	72	136	46	55	65	146
X position	237	180	141	163	282	455	326	235	412	113	266	257	363	158	423
Y position	182	332	388	188	374	296	332	432	488	46	302	23	22	81	95

Supplier			
Production quantity by period	826	X position	312
Inventory cost	0.3	Y position	363
Initial inventory	2042	Number of periods	5

This model includes three objective functions: inventory cost, transportation costs and CO_2 emissions. The result of the multiobjective model in a network conformed by one supplier and 15 customers, to be compared with the single transport optimization obtained by using the Vehicle Routing Problem – VRP. In the optimization processes two distinct genetic algorithms were used: a simple genetic algorithm for solving the VRP and an algorithm based on the NSGA2 for the multiobjective analysis, similar to what is presented in [28, 39, 40].

The emission factor of a typical urban goods distribution vehicle was used to analyze the CO_2 emission effect. The vehicle corresponds to a VAN with an average city emission of e = 190 g of CO_2 / km [41]. This parameter is multiplied by the number of kilometers traveled, in order to calculate the emitted CO_2 gases amount. The formulation for the multiobjective model with three objective functions is presented in Eq. 1 to 4.

$$minimizing\ G(g_1, g_2, g_3), \tag{1}$$

$$g_1 = \sum_{t\in\tau}\sum_{i\in v}\sum_{i\in v}\sum_{k\in K} c_{ij}x_{ij}^{kt}\ , \tag{2}$$

$$g_2 = \sum_{i\in v'}\sum_{t\in\tau} h_i I_i^t + \sum_{t\in\tau} h_0 I_0^t\ , \tag{3}$$

$$g3 = \sum_{t\in\tau}\sum_{i\in v}\sum_{i\in v}\sum_{k\in K} e\cdot c_{ij}\cdot x_{ij}^{kt}\ . \tag{4}$$

Equation (1) is the objective function that seeks to minimize transport costs, where x_{ij}^{kt} is a binary variable that is equal to 1 if the vehicle k has to travel from I to j in period t, and c_{ij} is the corresponding cost. Equation (2) is the functions for minimizing inventory both at the supplierI_0^t and at the customers I_i^t. Equation (3) minimizes the

CO2 emissions calculated as the sum of the distances multiplied by the vehicle CO2 emission factor. This objective function is restricted to the subsequent equations that assure the correct distribution process and correspond to those of the IRP Model according to Archetti et al., [42] and Arango et al. [6].

$$I_0^t = I_0^{t-1} + r_0^{t-1} - \sum_{k\in K}\sum_{i\in v} q_i^{kt-1} \;, \tag{5}$$

$$I_0^t \geq \sum_{k\in K}\sum_{i\in v} q_i^{kt}\, Y_i^{kt} \;, \tag{6}$$

$$I_i^t = I_i^{t-1} + \sum_{k\in K}\sum_{i\in v} q_i^{kt} - d_i^t \;, \tag{7}$$

$$I_i^t \geq 0 \;, \tag{8}$$

$$I_i^t \leq C_i \;, \tag{9}$$

$$q_i^{kt} \leq C_i - I_i^t \;, \tag{10}$$

$$q_i^{kt} \leq C_i Y_i^{kt} \;, \tag{11}$$

$$\sum_{i\in v} q_i^{kt} \leq Q_k \;, \tag{12}$$

$$\sum_{i\in v} q_i^{kt} \leq Q_k Y_0^{kt} \;, \tag{13}$$

$$\sum_{i\in v, i<j} X_{ij}^{kt} + \sum_{i\in v, j<i} X_{ji}^{kt} = 2y_i^{kt} \;, \tag{14}$$

$$\sum_{i\in S}\sum_{j\in S} X_{ij}^{kt} \leq \sum_{i\in S} y_i^{kt} - y_m^{kt} \quad \forall\, subset \;\; S \subseteq V \;, \tag{15}$$

$$q_i^{kt} \geq 0;\, Q_k \geq 0;\, I_i^t \geq 0;\, d_i^t \geq 0;\, C_i \geq 0. \tag{16}$$

For a thoughtful explanation of the restrictions, readers may refer to [6, 42] and [19]. The input parameters are obtained from the 15 customers and one supplier instance proposed by Archetti et al., [42]. In this case, the inventory amount that can be stored in each of the customers was increased, as a strategy to generate a lower distribution costs and a decreased CO_2 Emissions. The input data for location, initial quantity and inventory cost for customers and the supplier, as well as other instance information are presented in Table 1.

5 Results

The multiobjective model, in which the transport costs, inventory costs and CO_2 emission are optimized simultaneously, generates a set of 13 individuals due to the solutions non-dominance.

Table 2. Three objective functions multi-objective model results.

Solutions	Transport cost	CO_2 Emissions	Inventory cost	Total cost
1	3288.4	624.8	445.0	3733.3
2	4982.8	946.7	404.3	5387.1
3	5111.9	971.3	403.7	5515.6
4	3502.8	665.5	413.1	3915.9
5	5427.4	1031.2	403.7	5831.0
6	3902.6	741.5	406.9	4309.5
7	3642.3	692.0	407.5	4049.9
8	4156.1	789.7	404.9	4561.0
9	3304.3	627.8	421.3	3725.5
10	5012.8	952.4	404.2	5417.0
11	4831.7	918.0	404.8	5236.4
12	5692.9	1081.7	403.4	6096.3
13	3580.2	680.2	411.3	3991.5

Table 3. Results comparison for the models.

Model	Total Cost	Transport Cost	Inventory Cost	CO_2 Emissions
VRP.	8918.8	8612.5	306,3	1636,4
Multiobjective.	3733.3	3288.3	445.0	624.8

It is not possible to argue that one of the solutions is better than the others, for that reason, the decision maker, depending on his preference, may choose any of the model produced solutions.

Table 2 shows the results for the 3 optimized objective functions generated by the multobjective algorithm, presenting the CO_2 emissions, inventory, transport and total cost for each individual.

From Table 2 it can be observed that the lower the inventory level, the higher the transport cost and the CO_2 emissions, this as a consequence of an increase in transportation intensity in order to minimize inventory. This behavior is caused because the three-functions multiobjective model searches the best solution for every objective function without excessively increasing the others. In the solutions set produced by the multiobjective model, the individual number 1 is the solution that yields the lower CO_2 emissions, as observed in Table 1. Presented in Fig. 1, individual number 1, allows to serve the customers and satisfy its demand without visiting all customers in every period due to the collaboration between customers and supplier through the VMI. This generates an increase in inventory levels but reduces transportation and CO_2 emissions.

In order to compare these solutions, the distribution problem was solved supplying all customers in each period, what minimizes inventory costs in customers but increases transportation. For that, the Vehicle Routing Problem – VRP was used, assuming that

Inventory

Customer	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Period 1	0	167	0	0	0	45	0	0	109	0	0	215	275	0	427
Period 2	158	0	0	253	0	0	183	316	0	128	317	0	0	386	0
Period 3	0	0	288	0	329	0	271	0	0	0	0	0	0	0	0
Period 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Period 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Routes sequence (Routes starts on Depot -0- and ends on it.)

Period 1	0	6	9	15	12	13	2	0	0	0	0	0	0	0	0	0
Period 2	0	4	14	10	1	11	7	8	0	0	0	0	0	0	0	0
Period 3	0	5	3	7	0	0	0	0	0	0	0	0	0	0	0	0
Period 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Period 5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 1. Individual 1 generated by the three-functions multiobjective model.

every customer must be served every period, explaining why there is no customers inventory at the each period end, as well as no initial inventory need. For that distribution process, the transport cost is 1722.5 for each period, which corresponds to a total cost of 8612.5 for the 5 periods. This single VRP cost is higher than the transport cost for all individuals generated by the three-functions multiobjective model.

Table 3 presents the comparison of the distribution process with the VRP and individual 1 of the multobjective model, since this is the solution that generates the lower CO_2 Emissions. In Table 3 can be observe that the inventory is lower in the VRP solution, but it generates a higher costs and rises CO2 emissions, making it unattractive for companies and the environment.

The lower inventory cost in the VRP is caused by the transportation intensity that allows minimizing the inventory amounts required in customers facility. However, such intensity directly increases transport costs and the distances, what ultimately causes higher CO_2 emissions. A similar analysis could be made for the other solutions proposed by the algorithm and similar results will be found.

The results allow inferring that the multiobjective model generates solutions that improve the distribution process performance, both in cost and emissions, through evaluating different relationships between transport and inventory assignments. The proposed model results are better than the produced by the well-know VRP. However for a more comprehensive affirmation about the goodness of the multiobjective model, a comparison with more complex routing algorithms as well as trials with more and difficult instances are to be made.

6 Conclusions

In this paper, the collaborative inventory process and its effect on the pollutant gases emission are analyzed from a multiobjective perspective using an optimization model that includes three objective functions: inventory cost, transportation costs and CO_2 emissions, which are optimized simultaneously. This multiobjective model, as expected based on available literature, generates a set of optimal and non-dominated individuals, which achieves better results than the traditional, single transport optimization procedures, since total cost and CO_2 emissions are higher for the VRP. The increase in cost by the VRP model is caused by the inventory reduction at customers what mandates supplying customers every period. This behavior is known as a local optimum, in many cases worse than the global logistics optimization.

Through inventory collaboration it is possible to reduce goods distribution cost, and simultaneously minimize CO2 emissions due to the logistics activity. Based on the results found in this paper, the search for the single and individual optimization of transport or inventory costs generates large increases in logistical costs as well as CO2 emissions, which is neither beneficial for Company nor for the environment. However this conclusion applies only for the specific analyzed case. For a more comprehensive affirmation about the virtue of the multiobjective model, it should be tested with more complex instances, as well as compared with more sophisticated routing algorithms.

As future Work it will be interesting to study the possibility of including other objective functions that evaluates the performance of distribution processes such as Service level, process variability or risk [43]. It would also be interesting to consider distribution networks analysis involving several suppliers, several products and also more than one single Supplier-Customers echelons. Some authors have explored [6, 40, 43-44] these research lines, making their work an interesting starting point.

References

1. Arango-Serna, M. D., Adarme-Jaimes, W., Zapata-Cortes, J. A.: Inventarios colaborativos en la optimización de la cadena de suministros. Dyna, 80(181), pp. 71–80 (2013)
2. Visser, J., Nemoto, T., Browne, M.: Home Delivery and the Impacts on Urban Freight Transport: A Review. Procedia - Social and Behavioral Sciences 125, pp. 15–27 (2014)
3. Morana, J.: Sustainable Supply Chain Management in Urban Logistics. In: Gonzalez-Feliu, J., Semet, F., Routhier, J. L. (eds), Sustainable Urban Logistics: Concepts, Methods and Information Systems, EcoProduction (Environmental Issues in Logistics and Manufacturing), Springer, Berlin, Heidelberg (2014)
4. Zavanella, L., Zanoni, S.: A one-vendormulti-buyerintegratedproduction-inventorymodel: The ConsignmentStock case. International Journal of production Economics, 118, pp. 225–232 (2009)
5. Coelho, L. C., Cordeau, J. F., Laporte, G.: Consistency in multi-vehicle inventory-routing. Transportation Research Part C: Emerging Technologies, 24, pp. 270–287 (2012)
6. Arango, M. D., Zapata, J. A., Gutierrez, D.: Modeling: The Inventory Routing Problem (IRP) With Multiple Depots With Genetic Algorithms. IEEE Latin American Transactions. 13(12), pp. 3959 – 3965 (2015)
7. Chan, F. T. S., Prakash, A.: Inventory management in a lateral collaborative manufacturing supply chain: a simulation study. International Journal of production Research, 50(16), pp. 4670–4685 (2012)
8. Simatupang, T., Sridharan, R.: An integrative framework for supply chain collaboration. International journal of logistics management, 16, pp. 257–274 (2005)

9. Holweg, M., Disney, S., Holmström, J., Smaros, J.: Supply Chain Collaboration: Making sense of the strategy continuum. European Management Journal, 23(2), pp. 170–181 (2005)
10. Derroiche, R., Neubert, G., Bouras, A.: Supply chain management: a framework to characterize the collaborative strategies. International journal of computer integrated manufacturing, 21(4), pp. 426–439 (2008)
11. Díaz-Batista, J., Pérez-Armador, D.: Optimización de los niveles de inventario en una cadena de suministro. Ingeniería Industrial, 33(2), pp. 126–132 (2012)
12. Won-Cho, D., Hae-Lee, Y., Youn-Lee, T., Gen, M.: An adaptive genetic algorithm for the time dependent inventory routing problem. Journal of Intelligent Manufacturing, 25(5), pp. 1025–1042 (2014)
13. Bertazzi, L., Esperanza, M. G.: Inventory routing problems with multiple customers. EURO J Transp Logist 2, pp. 255–275 (2013)
14. Moin, N. H., Salhi, S., Aziz, N. A. B.: An efficient hybrid genetic algorithm for the multi-product multi-period inventory routing problem. Int. J. Production Economics, 133, pp. 334–343 (2011)
15. Rushton, P., Croucher-Baker, P.: The handbook of logistics and distribution management, 3rd edition. Ed. Kogan Page Limited (2010)
16. Estrada, M. A.: Análisis de estrategias eficientes en la logística de distribución de paquetería. Tesis Doctoral, Universitat Politècnica de Catalunya (2007)
17. Arango, M. D., Zapata, J. A., Adarme, W.: Aplicación del modelo de inventario manejado por el vendedor en una empresa del sector alimentario colombiano. Revista EIA, 15, pp. 21–32 (2011)
18. Gonzalez-Feliu, J., Peris-Pla, C., Rakotonarivo, D.: Simulation and optimization methods for logistics pooling in the outbound supply chain. Third International Conference on Value Chain Sustainability. Towards a Sustainable Development and Corporate Social Responsibility Strategies in the 21st Century Global Market, pp. 394–401 (2010)
19. Coelho, L. C., Laporte, G.: The exact solution of several classes of inventory-routing problems. Computers & Operations Research, pp. 558–565 (2013)
20. Zeng, Z., Zhao, J.: Study of stochastic demand inventory routing problem with soft time windows based on MDP. Adv. in Neural Network Research & Appli., LNEE (67), pp. 193–200 (2010)
21. Arango-Serna, M. D., Andrés-Romano, C., Zapata-Cortés, J. A.: Collaborative goods distribution using the IRP model. DYNA, 83(196), pp. 204–2012 (2016)
22. Campbell, A., Savelsbergh, M.: A Decomposition Approach for the Inventory-Routing Problem. Transportation Science, 38(4), pp. 488–502 (2004)
23. Seferlis, P., Pechlivanos, L.: Optimal Inventory and Pricing Policies for Supply Chain Networks. European Symposium on Computer-Aided Process Engineering (2004)
24. Arango, M. D., Zapata, J. A., Andres, C.: Metaheuristics for goods distribution. Proceedings of 2015 International Conference on Industrial Engineering and Systems Management (IESM), IEEE Publications, pp. 99–107 (2015)
25. Villalobos, M. A.: Análisis de Heurísticas de Optimización para Problemas Multiobjetivo. Tesis Doctoral, Departamento de Matemáticas, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional (2005)
26. López, J., Zapotecas, S., Coello, C.A.: An introduction to multiobjective optimization techniques.: in Ajith Abraham, Lakhmi Jain and Robert Goldberg (editors), Evolutionary Multiobjective Optimization: Theoretical Advances And Applications, pp. 7–32, Springer-Verlag (2009)
27. González-Álvarez, D.: Optimización Multiobjetivo y Paralelismo para Descubrir Motifs en Secuencias de AND. PhD. Thesis, Extremadura (2013)

28. Zapata-Cortes, J. A.: Optimización de la distribución de mercancías utilizando un modelo genético multiobjetivo de inventario colaborativo de m proveedores con n clientes. Tesis Doctoral, Universidad Nacional de Colombia.
29. Chen, C. L., Lee, W. C.: Multi-objective optimization of multi-echelon supply chain networks with uncertain product demands and prices". Computers and Chemical Engineering, 28, pp. 1131–1144 (2004)
30. Liang, T. F.: Fuzzy multi-objective production/distribution planning decisions with multi-product and multi-time period in a supply chain. Computers & Industrial Engineering, 55, pp. 676–694 (2008)
31. Liao, H. S., Hsieh, C. H., Lai, P. G.: An evolutionary approach for multi-objective optimization of the integrated location–inventory distribution network problem in vendor-managed inventory. Expert Systems with Applications, 38, pp. 6768–6776 (2011)
32. Azuma, R. M., Coelho, G. P., Von Zuben, F. J.: Evolutionary Multi-Objective Optimization for the Vendor-Managed Inventory Routing Problem. In EEE Congress on Evolutionary Computation (CEC), pp 1457–1464 (2011)
33. Azuma, R. M.: Otimização multiobjetivo em problema de estoque e roteamento gerenciados pelo fornecedor. Master Thesis, Faculdade de Engenharia Elétrica e de Computaçã. Universidade Estadual de Campinas (2011)
34. Shankar, B. L., Basavarajappa, S., Kadadevaramath, R. S., Chen, J. C. H.: A bi-objective optimization of supply chain design and distribution operations using non-dominated sorting algorithm: A case study. Expert Systems with Applications, 40, pp. 5730–5739 (2013)
35. Afshari, M., Sharafi, T., ElMekkawy, T., Peng, Q.: Optimizing multi- objective dynamic facility location decisions within green distribution network design. Procedia CIRP 17, pp. 675–679 (2014)
36. Nekooghadirli, N., Tavakkoli-Moghaddam, R., Ghezavati, V. R., Javanmard, S.: Solving a new bi-objective location-routing-inventory problem in a distribution network by meta-heuristics. Computers & Industrial Engineering, 76, pp. 204–221 (2014)
37. Pasandideh, S. H. R., Niaki, S. T. A., Asadi, K.: Optimizing a bi- objective multi-product multi-period three echelon supply chain network with warehouse reliability. Expert Systems with Applications, 42, pp. 2615–2623 (2015)
38. Pasandideh, S. H. R., Niaki, S. T. A., Asadi, K.: Bi-objective optimization of a multi-product multi-period three-echelon supply chain problem under uncertain environments: NSGA-II and NRGA. Information Sciences, 292, pp. 57–74 (2015)
39. Arango-Serna, M. D., Zapata-Cortes, J. A., Serna-Uran, C. A.: Collaborative Multiobjective Model for Urban Goods Distribution Optimization. In: García-Alcaraz J., Alor-Hernández G., Maldonado-Macías A., Sánchez-Ramírez C. (eds) New Perspectives on Applied Industrial Tools and Techniques. Management and Industrial Engineering. Springer, Cham (2018)
40. Arango, M. D., Zapata, J. A.: Multiobjective Model For The Simultaneous Optimization Of Transportation Costs, Inventory Costs And Service Level In Goods Distribution. EEE latin america transactions, 15(1), pp. 129–136 (2017)
41. Ford Motor Company. http://es.ford.com/trucks/ transitvanwagon/specifications/ (2016)
42. Archetti, C., Bertazzi, L., Laporte, G., Speranza, M. G.: A branch and cut algorithm for a vendor-managed inventory-routing problem. Transportation Science, 41(3), pp 382–91 (2007)
43. Fan, T., Chiang, W. C., Russell, R.: Modeling urban hazmat transportation with road closure consideration. Transportation Research Part D, 35, pp. 104–115 (2015)
44. Arango-Serna, M. D., Serna-Uran, C. A., Zapata-Cortes., J. A.: Multi-agent system modeling for the coordination of processes of distribution of goods using a memetic

algorithm. In: García-Alcaraz, J. G., Alor-Hernández, A., Maldonado-Macías, C., Sánchez-Ramírez (Eds), New Perspectives on Applied Industrial Tools and Techniques. Management and Industrial Engineering (2018)

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
marzo de 2018
Printing 500 / Edición 500 ejemplares