Advances in Decision Support Systems

Research in Computing Science

Series Editorial Board

Editors-in-Chief:

Grigori Sidorov (Mexico)
Gerhard Ritter (USA)
Jean Serra (France)
Ulises Cortés (Spain)

Associate Editors:

Jesús Angulo (France)
Jihad El-Sana (Israel)
Alexander Gelbukh (Mexico)
Ioannis Kakadiaris (USA)
Petros Maragos (Greece)
Julian Padget (UK)
Mateo Valero (Spain)

Editorial Coordination:

Alejandra Ramos Porras

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 132,** febrero 2017. 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 132**, February 2017. Printing 500. The authors are responsible for the contents of their articles. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission of Centre for Computing Research. Printed in Mexico City, in the IPN Graphic Workshop – Publication Office.

Volume 132

Advances in Decision Support Systems

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

Instituto Politécnico Nacional, Centro de Investigación en Computación
México 2017

ISSN: 1870-4069

Copyright © Instituto Politécnico Nacional 2017

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

The editors and the publisher of this journal have made their best effort in preparing this special issue, but make no warranty of any kind, expressed or implied, with regard to the information contained in this volume.

All rights reserved. No part of this publication may be reproduced, stored on a retrieval system or transmitted, in any form or by any means, including electronic, mechanical, photocopying, recording, or otherwise, without prior permission of the Instituto Politécnico Nacional, except for personal or classroom use provided that copies bear the full citation notice provided on the first page of each paper.

Indexed in LATINDEX, DBLP and Periodica

Printing: 500

Printed in Mexico

ISSN 1870-4069

Editorial

Within industrial contexts, managers are usually the top decision makers. They hold permanent contracts and have high incomes. Nowadays, advanced information and communication technologies support and assist these administrators in their decisions. They enable to consider several variables – including time costs – and perform calculations with greater precision. Such 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) computer vision and machine learning techniques to develop a counter medication classifier; 2) data mining methods to find relations and factors in medical opinions in a Mexican hospital about the decision of making autopsies 3) particle swarm to optimize the process of machining a workpiece, 4) the simulation model to analyze the sustainability of industrial ecosystems, 5) mapping tools and maximum coverage model to define a new of facility location of fire stations, 6) fuzzy QFD to analyze the operational risk assessment in 3PL for maritime transportation, 7) the economic order quantity (EOQ) model to reduce the inventory in a metal-mechanical industry and finally 8) a Bayesian predictive distribution and desirability function to optimize multiple response variables, 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 PRODEP.

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
February 2017

ISSN 1870-4069

Table of Contents

Page

ISSN 1870-4069

A Counter Medication Classifier Using Machine Learning and Computer Vision Techniques

Eduardo Ceh-Varela[1], Gandhi Hernández-Chan[1], Marisol Villanueva-Escalante[2], Giner Alor-Hernández[3], José Luis Sánchez-Cervantes[4]

[1] Universidad Tecnológica Metropolitana,
Information Technology and Communication School, Mérida,
Mexico

[2] Instituto Tecnológico de Mérida, Mérida,
Mexico

[3] Instituto Tecnológico de Orizaba, Division of Research and Postgraduate Studies, Orizaba,
Mexico

[4] Professor CONACYT commissioner at Instituto Tecnológico de Orizaba,
Mexico

{eduardo.ceh, gandhi.hernandez}@utmetropolitana.edu.mx,
{mvillanueva}@itmerida.mx, galor@itorizaba.edu.mx, jsanchezc@ito-depi.edu.mx

Abstract. Self-medication and self-prescription are common practices that can be observed in many countries around the world, from the most advanced in terms of medical services in Europe to the less ones as in South America or Africa. Self-medication is defined as the consumption of one or more drugs without the advice of a physician. Many studies in Mexico reveal the type of medications that are consumed as well the social groups that normally use this practice. The consequences for this can range from a mild allergic reaction to death. On the other hand, it is easy to buy drugs without a prescription in pharmacies or supermarkets, but consumers do not always know which one to choose, neither the ingredients nor side effects they can cause. Here we present a classifier model for counter medication based on computer vision and machine learning techniques. We collected 150 images from 11 different counter medications. The classifier was tested with 43 new images, and obtained 90.7% of accuracy, 93% of precision, 91% of recall and 91% of F1-score.

Keywords: Self-medication, computer vision, machine learning.

1 Introduction

Self-medication or self-prescription can be defined as the action of obtain and consume one or more medicines without the advice of a physician nor prescription for a treatment

[1, 2]. As it was stated by the World Health Organization (WHO) and the European Commission of Community Pharmacies, the use of drugs without prescription has been increased due the over-supply and the impact of the media communications concerning self-care to prevent and cure diseases [3, 4]. Studies like the one presented in [5] revealed that in México self–medication problems were observed mostly in women and in adult population with less access to health services. This study also states that the most consumed drugs are analgesics, antibiotics and antacids. Even when self-medication is a common practice, potential risks exist, for example an incorrect self-diagnosis, adverse reactions, incorrect administration or dosage, risk of dependence and abuse, delaying medical advice, and in the worst-case scenario to death [6, 7].

On the other hand, telematics applications have gained success in the field of health promotion, disease prevention and medical treatment. They are attractive because of the potential to provide information adapted to users' needs, social condition and access to technology. The benefits derived from the use of these applications can deeply affect the health status of the population, because they allow citizens to take better decisions regarding their health [8].

In this paper, we present a classifier model that uses machine learning and computer vision techniques to classify counter medications based on an image. We implement this classifier inside a web application that allows users to take a picture of a medication's box in order to receive information related to its content, such as its chemical composition, possible secondary effects, and general comments. The main motivation for this study was to present a tool that can be used to make better decisions about which counter medication to buy.

It is necessary to say that we do not intend to eliminate the problem of self-medication or encourage it, but we pretend to reduce the risks associated. We decided to work with counter medications which can be found in supermarkets, convenience stores or even drugstores.

In the next sections we present the state of the art about self-medication in México and artificial intelligence in the health field, the methodology that we followed for designing and building classifier, the results obtained from the different tests, and finally the conclusions.

2 Self-Medication in Mexico

In reference to Mexico, a study about self-medication in Morelia found that the most frequently purchased drugs were analgesics, anti-inflammatories, vitamins, antibiotics, flu, and benzodiazepines. The study revealed that 51.4% of sales were for self-medication and 13% for recommendation of the pharmacy clerk [9]. A survey applied to 1,537 individuals in Cuernavaca found that the 53.3% consumed drugs by self-medication, of these 64.9% were woman, and 7.6% were between 25 and 44 years old, and the most consumed drugs were analgesics and antibiotics [10].

A total of 245 residents of Tuxtla Gutierrez, in the state of Chiapas, participated in a study where more than half of the participants reported to have taken a medicine without prescription over the last 30 days. Self-medication was significantly more

frequent among older adults who lived alone, and among the illiterate or those with a low level of education [11]. Another survey applied to adults showed that self-medication is related to socioeconomic status and the lack of access to professional healthcare. It also states that the lack of government-sponsored health insurance coverage increases the propensity to self-medicate [12].

A survey applied to 1,859 university students in the state of Puebla showed that 96% of the respondents used drugs without medical advice. It also showed that 97.7% of the students have consumed AINES (acetylsalicylic acid, pseudoephedrine, analgesic, antipyretic) and its combinations, 42.8% mucolytics, 33.6% antidiarrheal, 32.9% antiparasitic, 28.8% antimicrobials, 6.5% topical drugs for acne treatment, and 5.8% oral contraceptives [4].

3 Artificial Intelligence in Medicine

This research turns around Artificial Intelligence (AI), especially in computer vision and machine learning. In this section we present some examples and cases where these areas have been used in medical and pharmaceutical fields.

The earliest work in medical artificial intelligence dates from the early 1970's in the United States, and now is a worldwide field with important contributions from around the globe [13]. Medicine is a field in which AI is critically needed because of the increasing expectations of healthcare [14]. In [15] it is presented a knowledge-based system to increase the overall effectiveness of physicians' time, and thus the quality of healthcare, by improving the information exchange between physicians and patients in clinical settings. The research in [16] focuses on the use of machine learning and cluster analysis techniques to classify a set of organic molecules into their pharmacological activity of sedative and tranquilizer. This was based on the premise that a pharmacological activity of any molecule is dependent on its structure and that structural changes may lead to changes in the activity.

The use of drugs is also a field addressed by AI. Several programs focusing on drug's therapy have been developed. They provide guidance on drug interactions, drug therapy monitoring, and drug formulary selection. There are many aspects of pharmacy where AI can have an impact on [17].

Computer vision is part of the AI field which is centered in theoretical studies of visual information processing. Its two main goals are to develop image understanding systems, and to understand human vision [18]. From its beginning, computer's graphics have been applied to medical fields with research challenges. In modern medicine, medical imaging systems are only one prominent example [19]. In this field, one of the used techniques is known as image matching, which can be considered as a classification scheme than can be used with 2D and 3D images and other signal modalities as electroencephalography and magnetoencephalography [20].

In medicine, the analysis of 3D images can improve significantly diagnosis and therapy [21]. The multimedia workstation of a physician could not exist without tools for manipulating images to perform measurements and to extract and collect

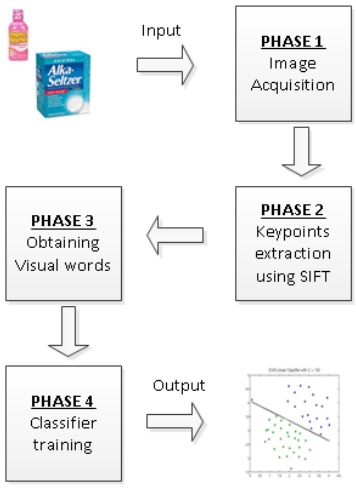


Fig. 1. Four phase methodology.

information. Image analysis and computer vision constitute a wide and rapidly evolving field [22].

Machine learning is used to extract information from the raw data in databases. The process consists on taking the data, and inferring whatever structure in it [23]. Today machine learning provides several indispensable tools for intelligent data analysis, and is currently well suited for analyzing medical data, for medical diagnosis [24]. Another example of the use of machine learning in medical field is the framework proposed by [25]. It is a content-based image retrieval framework for a collection of medical images, and consists of machine learning methods for image prefiltering, similarity matching using statistical distance measures, and a feedback scheme

4 Methodology

We extended the methodology presented in [26] to a four phase methodology as can be seen in figure 1. The first phase was the acquisition of images of counter medications that served to train the classifier. In the second phase the SIFT algorithm was used to extract the images relevant characteristics as keypoints. In the third phase, we used the visual-bag-of-words method to extract a visual word vector for each image; this information was added to the data set to be used in the next step. Finally, the data set was used to train a classifier in order to create a classifier model that identifies new instances of counter medications.

4.1 Image Acquisition

We collected 150 images from 11 different counter medications found on supermarkets and drug stores. The images were collected using the camera of an iPhone 6 and a Sony

Xperia Z2; also we used images downloaded from Internet. Table 1 shows the commercial name for each medication and the number of collected images.

The images were taken with the cell phone cameras in different angles, backgrounds, and lighting. They were JPEG and PNG file format. Also, we group together different boxes of the same medication. An example of these images can be seen in figure 2. We also resized all the images to a same size because of the different images' resolution and convert all to PNG format for better handling.

Table 1. Medication images distribution.

#	Commercial name	Number of images
1	Alka seltzer	11
2	Aspirin	11
3	Bengay	11
4	Bepanthol	11
5	Iliadin	11
6	Pepto bismol	27
7	Sedalmerk	9
8	Sensibit D	11
9	Tempra	19
10	Treda	11
11	Tylenol	18
Total		150

Fig. 2. An example of the images used for training.

4.2 Features Extraction

We use the Scale Invariant Feature Transform (SIFT) method to detect local invariant descriptors inside of each of the medication's images. One of the advantages of using SIFT is that descriptors do not change if the image is scaled, rotated, or even its intensity is changed [31]. It uses the difference-of-Gaussian (DOG) functions to identify potential interest points (keypoints), where these points across both image location and scale are the maxima and minima of (1).

$$D(\boldsymbol{x}, \sigma) = [G_{k\sigma}(x) - G_{\sigma}(x)] * I(x) = [G_{k\sigma} - G_{\sigma}] * I = I_{k\sigma} - I_{\sigma} \qquad (1)$$

This method computes a descriptor for each keypoint based on its position, scale and rotation [32]. For each keypoint a feature vector was created. This vector is formed by a 4x4 keypoint subregions with 8 bins orientation histogram, resulting in a 128 bins histogram (4x4x8).

In order to apply this algorithm to our project we used the library OpenCV 2.4.9. On the left side of figure 3 the keypoints are shown for a medication box, and on the right side is the SIFT histogram for one of those keypoints.

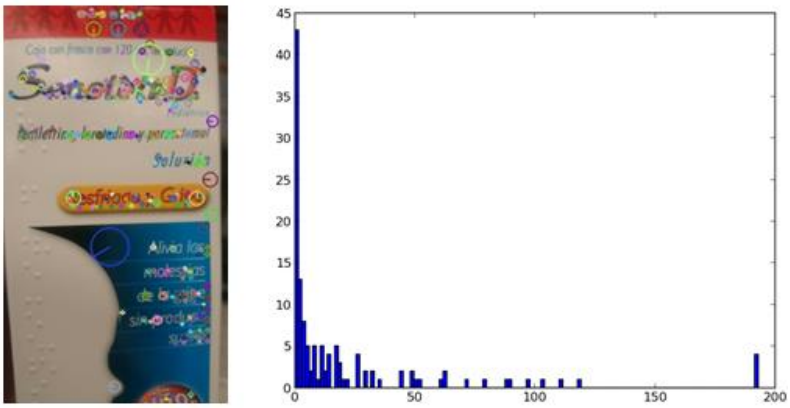


Fig. 3. A SIFT orientation histogram for one of the Sensibit D box keypoints on the left.

4.3 Dataset Creation

To create our dataset we applied the concept of the Bag-of-Words (BOW) method to the image's vectors extracted from the keypoints. The Bag-of-Words method is commonly used to classify text documents, where each text document can be represented by the frequency with which each word appears inside the document [27]. The process of the visual-word vectors' extraction is presented in the figure 4.

In order to use this method, the descriptor vectors extracted on the previous phase were grouped into clusters of similar descriptors. Each of these clusters represents a "visual word", in this manner an image can be represented as a Bag-of-visual-words (BOVW), then a visual vocabulary was created. Thus, for each image a vector is

generated containing the histogram of each visual-word. Then, these vectors can be used as features in a classifier [29-31].

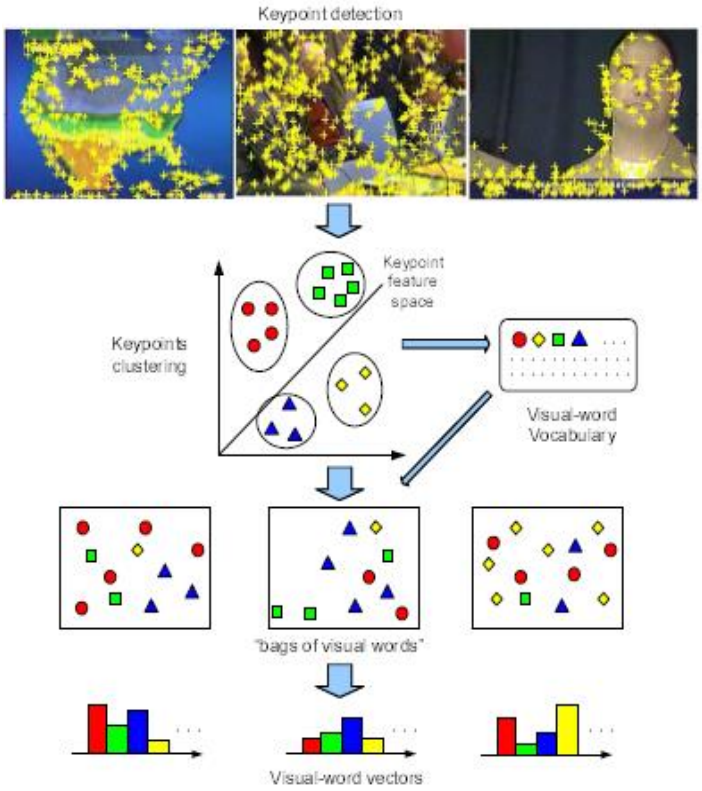


Fig. 4. Extracting the visual-word vectors [28].

For our project we used the K-means algorithm from the Scipy library to form clusters. This algorithm uses as parameter the number of clusters (*k*) to find. To determine the right *k* number of clusters we tried two of the methods presented in [32]. The first method we used was the rule of thumb in (2) for estimating the k number of clusters. This method gave us a value of $k \approx 143$.

$$k \approx \sqrt{n/2} \tag{2}$$

The second method we used was the Elbow Method. This is a visual method to estimate the *k* number of clusters. The main idea of this method is to continuously increment the value of *k* until the cost function changes in a dramatic way, and then it reaches a plateau. However, sometimes the change cannot always be unambiguously identified. We used a range of *k* values from 50 through 450. As figure 5 shows using this method at the value of *k*=150 we observe the change.

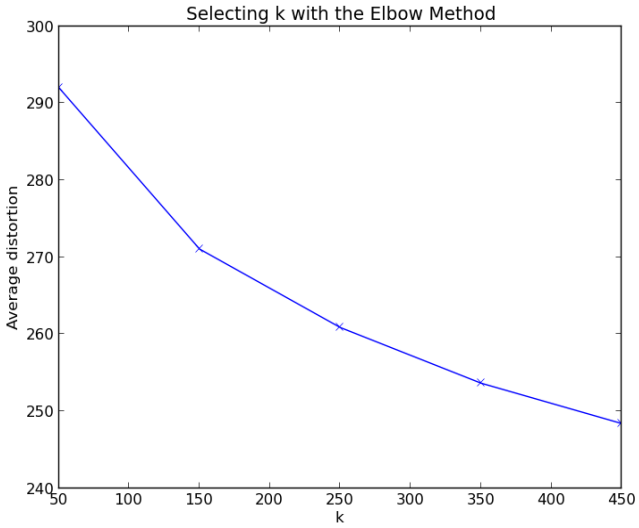


Fig. 5. The elbow plot to choose *k*.

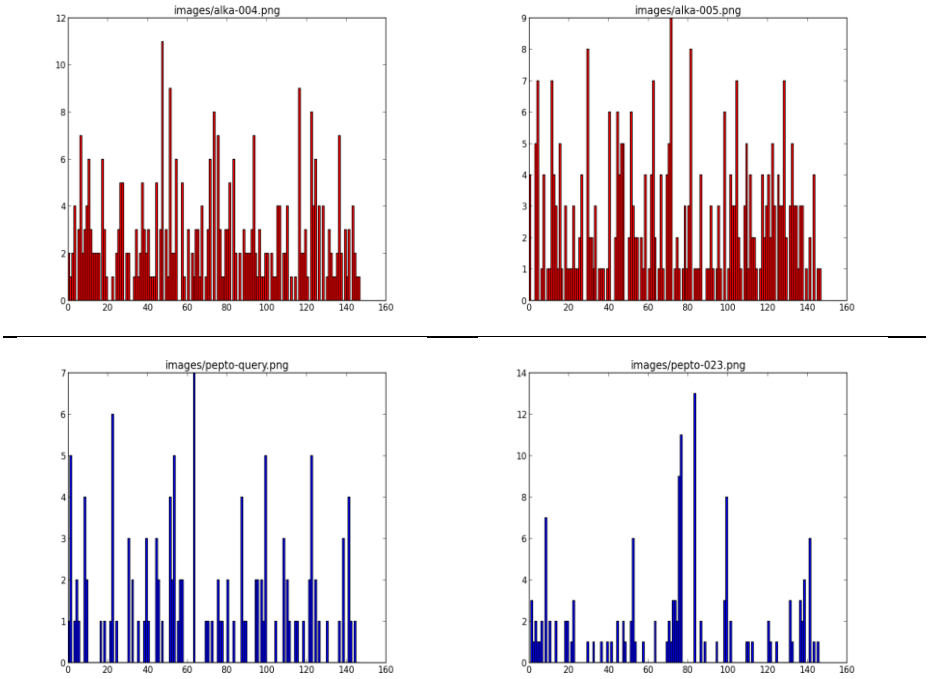


Fig. 6. Visual word histograms. At the upper row are histograms for Alka Seltzer boxes, and at the lower row are histograms for different Pepto Bismol bottles.

As both obtained values were close together, we decided to use their average as the number of *k*. Using $k \approx 147$, the K-means algorithm returned the centroids for each cluster, which means that we ended up with a vocabulary of 147 visual words.

Once we had this vocabulary, we extracted the frequency of visual words that exists for each image in order to obtain a histogram vector for each of them. In figure 6 we can visualize the histogram for four different medication boxes. The upper row contains the histograms for two different Alka Seltzer boxes, and at the bottom row are the histograms for two different Pepto Bismol bottles. It can be seen the different histogram shapes for the two different counter medications and the similarity between the same medications.

4.4 Classifier Model Generation

We split the dataset in a train set with the 70% of the images (n=107), and the remaining images formed the test set (n=43). A stratified sampling was used with the intention that all the classes were equally distributed between the two datasets.

To generate the model we used Support Vector Machines (SVM) [33] algorithm. This is a supervised learning algorithm that has grown in popularity for classification and regression due to its excellent results [34, 35]. One of its advantages is the effectiveness when working with high-dimensional data and little number of samples [36]. Based on a labeled training data, this algorithm tries to find an optimal hyperplane that maximizes the separation margin between different classes. We used SVC implementation from the Scikit-Learn machine learning library in Python.

One important aspect to consider when using the SVM algorithm is to choose the right parameters to build the classifier. We test several values in a grid search strategy to obtain the best configuration parameters C and gamma, and also for the type of Kernel to use. The best parameters were for C=10, gamma=0.001, kernel='rbf'.

Training Confusion Matrix

```
[[ 7  0  0  0  0  1  0  0  0  0  0]
 [ 0  6  0  0  0  0  0  0  0  0  0]
 [ 0  0 19  0  0  0  0  0  0  0  0]
 [ 0  0  0  8  0  0  0  0  0  0  0]
 [ 0  0  0  0 13  0  0  0  0  0  0]
 [ 0  0  0  0  0  8  0  0  0  0  0]
 [ 0  0  0  0  0  0  8  0  0  0  0]
 [ 0  0  0  0  0  0  1  7  0  0  0]
 [ 0  0  0  0  0  0  0  0 13  0  0]
 [ 0  0  0  0  0  0  0  0  0  8  0]
 [ 0  0  0  0  0  1  0  0  0  0  7]]
```

Training Classification Report

	precision	recall	f1-score	support
0	1.00	0.88	0.93	8
1	1.00	1.00	1.00	6
2	1.00	1.00	1.00	19
3	1.00	1.00	1.00	8
4	1.00	1.00	1.00	13
5	0.80	1.00	0.89	8
6	0.89	1.00	0.94	8
7	1.00	0.88	0.93	8
8	1.00	1.00	1.00	13
9	1.00	1.00	1.00	8
10	1.00	0.88	0.93	8
avg / total	0.98	0.97	0.97	107

Fig. 7. Confusion matrix and classification report for the training dataset.

The SVM classifier for the training dataset had a 97.1% of accuracy, with an average of 98% for precision and 97% for recall. In figure 7 the confusion matrix for the training dataset is shown on the left, and the classification report is on the right. Precision is the ratio of how correct is our classifier when it says that an instance is positive. Similarly, the result of Recall is the proportion of positive cases that are classified correctly [36].

For its part, the F1-score is the harmonic mean that attempts to combine the two previous measures [36, 37]. In the same report it can be seen the result of the confusion matrix for the training data set. In this matrix, the diagonal elements represent the number of correctly classified elements; meanwhile those off-diagonal elements represent the number of incorrectly classified elements. The greater the number of elements found in the diagonal of the matrix, the better the classifier model.

5 Results

The resulting classifier was tested with the test dataset, which consisted on 43 images of counter medications. The accuracy for the classifier was of 90.7%, the confusion matrix for this test can be seen in figure 8. In addition, for this test the precision was of 93%, the recall of 91%, and the F1-score was of 91%.

Test Confusion Matrix

```
[[2 0 0 0 0 0 1 0 0 0 0]
 [0 2 0 0 0 0 1 0 0 0 0]
 [0 0 8 0 0 0 0 0 0 0 0]
 [0 0 0 3 0 0 0 0 0 0 0]
 [0 0 1 0 5 0 0 0 0 0 0]
 [0 0 0 0 0 3 0 0 0 0 0]
 [0 0 0 0 0 0 3 0 0 0 0]
 [0 0 0 0 0 0 0 3 0 0 0]
 [0 0 0 0 0 0 0 0 5 0 0]
 [0 0 0 0 0 0 0 0 0 3 0]
 [0 0 1 0 0 0 0 0 0 0 2]]
```

Test Classification Report

	precision	recall	f1-score	support
0	1.00	0.67	0.80	3
1	1.00	0.67	0.80	3
2	0.80	1.00	0.89	8
3	1.00	1.00	1.00	3
4	1.00	0.83	0.91	6
5	1.00	1.00	1.00	3
6	0.60	1.00	0.75	3
7	1.00	1.00	1.00	3
8	1.00	1.00	1.00	5
9	1.00	1.00	1.00	3
10	1.00	0.67	0.80	3
avg / total	0.93	0.91	0.91	43

Fig. 8. Confusion matrix and classification report for the test dataset.

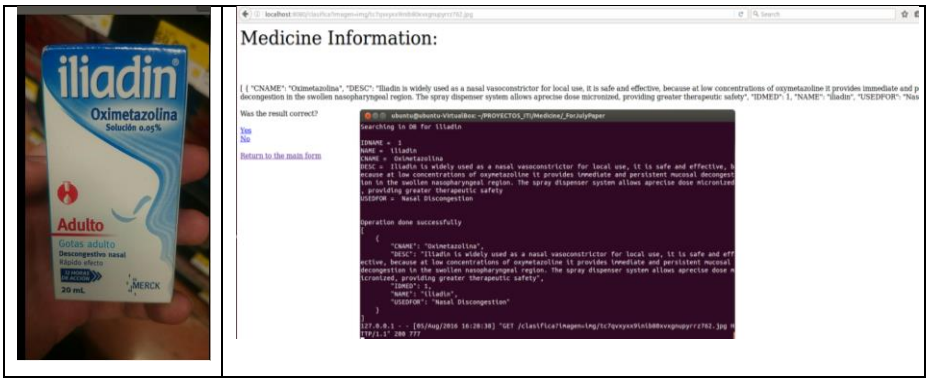


Fig. 9. New image classification and information retrieval.

To use the classifier model we developed a simple web application. From this application the user can upload the image of a counter medication and then ask the classifier to indicate the type of medication. Additionally, we created a database with

information about each medication. The result provided by the classifier is the input to query the database in order to retrieve the information related to the counter medication.

An example can be seen on the figure 9, where we used an image of the medicine Iliadin (Oximetazolina) taken directly from the smartphone's camera at a supermarket and uploaded to the web application though a web form. This image was not included on the initial dataset used to train or test the classifier. The uploaded image served as an input for the classifier and its output was used to query the database. As can be seen, the system predicted the correct medicine's class and returned the information regarding its name, compound, description and for what it is used. This information is returned in a JSON format to be used in other applications. The web application asks for the user feedback, in order to obtain more images for train the classifier seeking for accuracy improvement.

6 Conclusions

Because counter medication is easy to acquire in supermarkets, or convenience stores self-medication has become a common practice in many countries. Many studies in Mexico have shown that people use medication without prescription. In some cases this could have potential risks and consequences.

Machine learning and computer vision techniques are becoming a trend in healthcare applications in order to improve people's health. In this study we presented a classifier model aimed to help people who use counter medication to make better decisions. Our classifier uses SIFT and Visual-Bag-Of-Words techniques to extract features from counter medication boxes. With these features we generated a dataset that was used to train a SVM classifier. Our classifier got an accuracy of 90.7%, with a precision of 93%, a recall of 91%, and an F-1 score of 91% for the test dataset.

For testing purposes, the classifier was integrated in a web application to be used for people who wants to have more information about the counter medication he/she pretends to buy. In the same manner, this web application will help us to collect more images for training in order to improve the classifier's accuracy.

As a future work we want to expand the web application to include a recommendation system based on sentiment analysis, which uses our classifier to increase the information to be shown to users. This additional information will be extracted from Medic-Us, which is a social network for physicians and patients that allow the communication among them [38]. We also want to increment the number of counter medication classes known by the classifier, as well the number of images in dataset.

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

References

1. Peach, H.: Trends in self-prescribing and attitudes to self-medication. The Practitioner, 227(1384), pp. 1609 (1983)
2. Montastruc, J. L., Bagheri, H., Geraud, T., Lapeyre-Mestre, M.: Pharmacovigilance of self-medication. Therapie, 52(2), pp. 105–110 (1996)
3. Sanz, F., Loza, M. I., Ahlgrimm, E. D., Baetens, P., Cuypers, J., Cranz, H., Sosa-Iudicissa, M.: Telematics in community pharmacies to support responsible self-medication. The TESEMED project, Studies in health technology and informatics, pp. 141–148 (1996)
4. de Celis, E.S.P., Nava, Y. R.: Patrones de autoatención y automedicación entre la población estudiantil universitaria de la ciudad de Puebla. Elementos: ciencia y cultura, 11(56), pp. 43–51 (2004)
5. Ch, A.A.: Estudio de Automedicación en una farmacia comunitaria de la ciudad de Toluca. Revista Mexicana de Ciencias Farmacéuticas, pp. 5–11 (2009)
6. Ruiz, M.E.: Risks of self-medication practices. Current drug safety, 5(4), pp. 315–323 (2010)
7. Hughes, C.M., McElnay, J. C., Fleming, G. F.: Benefits and Risks of Self Medication. Drug safety, 24(14), pp. 1027–1037 (2001)
8. Alonso, A., Diaz, C., Sanz, F.: A matrix-approach proposal for the treatment of contents to be integrated into multimedia health resources for citizens. Studies in health technology and informatics, 68, pp. 3–6 (1998)
9. Acevedo, D.V., Alfaro Valle, A., Martínez Toledo, J. L.: Características de la adquisición de medicamentos en Morelia (Michoacán, México). Available from: http://iris.paho.org/xmlui/handle/123456789/15554 (1995)
10. Angeles-Chimal, P., Medina-Flores, M. L., Molina-Rodríguez, J. F.: Automedicación en población urbana de Cuernavaca, Morelos. Salud pública de México, 34(5), pp. 554–561 (1992)
11. Balbuena, F.R., Aranda, A. B., Figueras, A.: Self-Medication in Older Urban Mexicans. Drugs & Aging, 26(1), pp. 51–60 (2009)
12. Pagán, J.A., Ross, S., Yau, J., Polsky, D.: Self-medication and health insurance coverage in Mexico. Health policy, 75(2), pp. 170–177 (2006)
13. Patel, V.L., Shortliffe, E. H., Stefanelli, M., Szolovits, P., Berthold, M. R., Bellazzi, R., Abu-Hanna, A.: The coming of age of artificial intelligence in medicine. Artificial intelligence in medicine, 46(1), pp. 5–17 (2009)
14. Szolovits, P.: Artificial Intelligence in Medicine. Boulder, CO, Westview Press (1982)
15. Buchanan, B.G., Moore, J. D., Forsythe, D. E., Carenini, G., Ohlsson, S., Banks, G.: An intelligent interactive system for delivering individualized information to patients. Artificial intelligence in medicine, 7(2), pp. 117–154 (1995)
16. Chu, K.C.: Applications of artificial intelligence to chemistry. Use of pattern recognition and cluster analysis to determine the pharmacological activity of some organic compounds. Analytical chemistry, 46(9), pp. 1181–1187 (1974)
17. Dasta, J.F.: Application of artificial intelligence to pharmacy and medicine. Hospital pharmacy, 27(4), pp. 312–315 (1992)

18. Poggio, T., Torre, V., Koch, C.: Computacional vision and regularization theory. Nature, 317(6035), pp. 314–319 (1985)
19. Gross, M.H.: Computer Graphics in Medicine: From Visualization to Surgery Simulation. ACM SIGGRAPH Computer Graphics, 32(1), pp. 53–56 (1998)
20. Van den Elsen, P.A., Pol, E. J., Viergever, M. A.: Medical image matching-a review with classification. Engineering in Medicine and Biology Magazine, pp. 26–39 (1993)
21. Ayache, N.: Medical computer vision, virtual reality and robotics. Image and Vision Computing, 13(4), pp. 295–313 (1995)
22. Pun, T., Gerig, G., Ratib, O.: Image analysis and computer vision in medicine. Computerized Medical Imaging and Graphics, 18(2), pp. 85–96 (1994)
23. Witten, I.H., Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann (2005)
24. Kononenko, I.: Machine learning for medical diagnosis: history, state of the art and perspective. Artificial Intelligence in medicine, 23(1), pp. 89–109 (2001)
25. Rahman, M.M., Bhattacharya, P., Desai, B. C.: A Framework for Medical Image Retrieval Using Machine Learning and Statistical Similarity Matching Techniques With Relevance Feedback. IEEE transactions on Information Technology in Biomedicine, 11(1), pp. 58–69 (2007)
26. Ceh-Varela, E., Hernandez-Chan, G.: Tomatoes classifier using color histograms. In: International Conference on Advances in Engineering Science and Management (2015)
27. Layton, R., Learning Data Mining with Python. Packt Publishing Ltd. (2015)
28. Yang, J., Jiang, Y. G., Hauptmann, A. G., Ngo, C. W.: Evaluating bag-of-visual-words representations in scene classification. In: International workshop on multimedia information retrieval (2007)
29. Filliat, D.: A visual bag of words method for interactive qualitative localization and mapping. In: Proceedings IEEE International Conference on Robotics and Automation, IEEE (2007)
30. Yang, J., Jiang, Y. G., Hauptmann, A. G., Ngo, C. W.: Evaluating bag-of-visual-words representations in scene classification. In: Proceedings of the international workshop on Workshop on multimedia information retrieval, ACM (2007)
31. Nicosevici, T., Garcia, R.: Automatic visual bag-of-words for online robot navigation and mapping. IEEE Transactions on Robotics, 28(4), pp. 886–898 (2012)
32. Kodinariya, T.M., Makwana, P. R.: Review on determining number of Cluster in K-Means Clustering. International Journal of Advance Research in Computer Science and Management Studies, 1(6), pp. 90–95 (2013)
33. Harrington, P.: Machine Learning in Action. Vol. 5, Manning Publications Co. (2012)
34. Cherkassky, V., Ma, Y.: Practical selection of SVM parameters and noise estimation for SVM regression. Neural networks, 17(1), pp. 113–126 (2004)
35. Moreno, P.J., Ho, P. P., Vasconcelos, N.: A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications. In: Advances in neural information processing systems (2003)

36. Garreta, R., Moncecchi, G.: Learning scikit-learn: Machine Learning in Python. Packt Publishing Ltd. (2013)
37. Powers, D.M.: Evaluation: from Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation. Journal of Machine Learning Technologies, 2(1), pp. 37–6 (2011)
38. Hernández-Chan, G., Ceh-Varela, E., Jimenez-Kantun, M.: Medic-Us Red Social Inteligente. In: Avances y Perspectivas de la Innovación y Vinculación, Universidad Tecnológica Metropolitana, pp. 65–76 (2016)

ISSN 1870-4069

Preliminary Results of an Analysis Using Association Rules to Find Relations between Medical Opinions About the non-Realization of Autopsies in a Mexican Hospital

Elayne Rubio Delgado[1], Lisbeth Rodríguez-Mazahua[1], Silvestre Gustavo Peláez-Camarena[1], María Antonieta Abud-Figueroa[1], José Antonio Palet Guzmán[2], Asdrúbal López-Chau[3], Giner Alor-Hernández[1]

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

[2] Hospital Regional de Rio Blanco (H.R.R.B.), Veracruz, Mexico

[3] Universidad Autónoma del Estado de Mexico, Zumpango, Estado de Mexico, Mexico

{lrodriguez, galor}@itorizaba.edu.mx, {gpelaez, mabud}@ito-depi.edu.mx, n3nyrbd@gmail.com, palet_guzman@hotmail.com, alchau@uaemex.mx

Abstract. In the last years, a significant reduction in the number of autopsies realized in the hospitals of the world has been observed. Since medics are the closest people to this problematic, they can offer information that helps clarify why the decreasing of this practice has occurred. In this paper, data mining techniques are applied to perform an analysis of medical opinions regarding the realization of autopsies in a hospital of Veracruz, in Mexico. The opinions were collected through surveys applied to 85 medics of the hospital. The result is a model represented by a set of rules that suggests some of the factors that are related to the decrease in the number of autopsies in the hospital, according to the survey respondents.

1 Introduction

Autopsies are one of the most important exams of the medical practice. They allow obtaining vital information about diseases, their development and the effects in the human organism. Nevertheless, the realization of autopsies supposes problematic and interesting situations of acceptance or rejection in the population [1].

In the last years, a decrease in the practice of autopsies in a worldwide level has been observed [1]. In the hospital centers of the biggest cities of Mexico, and particularly in the state of Veracruz, this decreasing trend in the number of autopsies performed is

presented. However, to the best of our knowledge, there are not published studies about the causes related to this phenomenon.

In this sense, it is pertinent to directly and specifically investigate the motives for the non-realization of autopsies in the hospitals, focusing the attention on the medics and authorities trained to act on that subject.

In this paper, the preliminary results of an analysis performed to datasets obtained by the application of an instrument (survey) to medics of a hospital of Veracruz are presented. Association rules were used in such analysis because they allow finding relations between medical opinions regarding several causes for the non-realization of autopsies, as well as the level of studies, years of experience, and participation of the medics in autopsy cases, to mention a few.

For a better explanation of the object of study, the rest of the paper is structured as follows. Section 2 shows the relating to the state of the art and Section 3 describes the case of study. Finally, the conclusions and the future work are given in Section 4.

2 Related Works

With the motivation of supporting the development of the health sector in smart cities, Oviedo Carrascal, et al. [2] described a set of advances and trends of data mining in this area. Nowadays, it is known that several medicines produce adverse effects and the efficient identification of these drugs is a challenge for the experts. Wang, et al. [3] detected relations between diseases and illnesses using association rules derived from the clinical evidence and analysis of cases. The factors that influence the survival to the disease of cervix cancer and the time of survival to the disease were identified in [4] through the application of decision trees and association rules. The study presented in [5] demonstrated that by means of the application of association rules to a dataset of patient records, it is possible to obtain correlations between the different attributes, such as exams, medicines, treatments and patient profiles. In [6] it was presented the new prediction system ICU ARM-II (Intensive Care Unit Association Rule Mining), based in a set of association rules that forms a flexible model for the prediction of personalized risk. This approach assumes a classification supported by association. The risk factors correlated to Diabetes Mellitus type 2 (DM2) were identified in [7] using association rules.

Embracing other areas, in [8] the effectiveness of online collaborative learning environment and the predictors generated using data mining (DM) techniques were evaluated. The study [9] examined the pattern of student performance by using the "K-means" clustering technique. In the study [10] authors offered a general overview of the recent works undertaking the predictive and descriptive tasks in the building field, with the aim of improving building performance. On the other hand, the works [11, 12] confirmed the considerable potential of Bayesian networks to explore coastal databases. The software industry has also benefited from DM techniques as seen in [13], where Ahmed H. Yousef proposed a solution architecture that enhances software development based on data in software repositories and presented a benchmark that provides an ensemble of DM models in the defective modules prediction problem and compares the

results. In [14] an automatic real-time web usage data mining and recommendation system is presented. The system is powered by K-NN (K-Nearest Neighbor) classification model implemented with Euclidian distance method. In [15] DM methods were analyzed for detection of anomalies in social networks. The research focuses on the different types of anomalies, classification based on features along with the assumptions and the reasons for their existence, as well as techniques to prevent and detect them. The work [16] described the procedure to recognize patterns that allow evaluating the energy efficiency of buildings by applying the K-means algorithm.

Table 1. Related works.

Work	DM techniques	Area	DM Tasks
[2]	Decision trees, Neural networks, "K-means"	Health	Classification, Clustering
[3]	Association rules	Health (Pharmacology)	Association
[4]	Decision trees and association rules	Health	Classification, Association
[5]	Association rules	Health	Association
[6]	CBA (Classification based in Association)	Health	Classification, Association
[7]	Association rules	Health	Association
[8]	Decision trees	Education	Classification
[9]	Clustering based in partition	Education	Clustering
[10]	Support Vector Machines. Neural network Decision trees Association rules	Engineering	Classification, Clustering, Association
[11]	Bayesian methods	Ecology	Classification
[12]	Bayesian methods	Ecology	Classification
[13]	Neural network Decision trees Naive Bayes	Software	Classification
[14]	K-Nearest Neighbor	Internet: Online applications	Classification
[15]	K-Nearest Neighbor Support Vector Machines. Neural network	Social networks	Classification Clustering,
[16]	Decision trees	Engineering	Classification

As we can see in Table 1, the different studies mentioned demonstrate the usefulness of data mining techniques for the solution of problems in several areas of the modern reality, and the medical area is raised as a great object of study, with appropriate problems to be studied from this perspective [2-7]. Nevertheless, to the best of our

knowledge there are not works which used association rule mining to analyze the decrease in the number of autopsies performed in a hospital, therefore this determines the appropriateness, novelty and interest of this research.

3 Analysis of Relations between Medical Opinions

3.1 Association Rules

This research is aimed to perform a descriptive analysis through DM, which allows the identification of the relations existing in the survey data and obtaining a model based on rules. With this model it is possible to evaluate which are the factors that influence in the decreasing of the practice of autopsies. For this reason, association rules were considered for this study.

Association rules [17] allow the discovering of relations between attributes based on the occurrence frequency of the attributes in transactions. An association rule is an implication of X over Y representing the form $X \Rightarrow Y$, where X and Y are sets of disjunct elements and the rule is interpreted as a norm that when X appears, Y also tends to appear. These rules are evaluated with the measures of support, which represents the ratio of elements of the dataset that contains X and Y, and confidence, which is the probability that a transaction containing X also contains Y. The association algorithms perform automatic searches of rules that relate sets of attributes and evaluate whether those rules are statistically significant. To do this, Weka basically has the Apriori [18], FilteredAssociator [19], FPGrowth [20], GeneralizedSequentialPatterns [21], PredictiveApriori [22] and Tertius [23] algorithms.

Apriori [24] is a classic algorithm for association rule mining. Apriori generates rules by means of an incremental process that performs searches of frequent relations between attributes delimited by a minimum confidence. The algorithm can be configured to be executed under certain criteria such as upper and lower limits of coverage to accept itemsets that accomplish the minimum confidence constraint, which is a ranking criteria to show the rules, and a parameter to indicate the specific quantity of rules that we want to find. Apriori it only seeks rules between symbolic nominal attributes and generates better results than other algorithms mentioned above, as it is shown in [25].

3.2 Dataset Collection and Processing

With the objective of analyzing the possible causes of the decrease in autopsies performed in the hospital system of "Servicios de Salud de Veracruz", a survey elaborated by one of the authors of this work was used as social research technique. Such survey was applied to 85 medics of one of the hospitals of the above mentioned system. The survey fundamentally pretended to investigate the opinion of the medics facing the practice of autopsies, as well as the principles, beliefs or motives that characterize the medics. The answers of the medics were transformed to a suitable representation to apply the data mining techniques in order to significantly reduce the

complexity of the data analysis. Table 2 shows a summary of the applied survey, and the amount of categories generated by each question. Figure 1 presents the percentage of questions that belongs to each aspect explored by the survey.

The answers of the surveys applied were introduced in a binary matrix, in this matrix every row represents a survey respondent, the columns represent every answer. The value of each column is the intersection that can be read as a pair <*answer*, *value*>, where *value* is equal to 'Y' if the survey respondent gives such answer and 'N' otherwise. Such dataset is denominated in this article as **C**.

Table 2. Summary of the survey applied to the medical staff.

Aspects	**Questions**	**Type of question**	**Generated categories**
Medical training	Area	Closed	3
	Grade	Closed	3
	General Medicine Training Center	Closed	47
	Medical Specialty Training Center	Closed	47
Medical experience	Years of medical practice	Closed	5
	Participation in Autopsy cases	Closed	5
Medical opinion about the discoveries in autopsies	Cause discrepancy with the clinical diagnoses	Closed	5
	Originate in claim cases	Closed	5
	Originate in arbitration cases	Closed	5
Survey respondent opinion about the request of autopsies	Motives for autopsy acceptance	Open	26
	Motives for autopsy rejection	Open	27
	Motives for autopsy rejection by family	Closed	9
	Motives for not enough autopsies performed in the hospital	Closed	8
Medical opinion about the procedure to request an autopsy	Suitable staff to request an autopsy	Closed	6
	Efficient methods to request an autopsy	Closed	14
General aspect	Comments	Open	25

With the objective of analyzing relations between the medical experience and the medical opinion about the procedure of autopsy requests, a subset $\mathbf{C_1}$ was generated from **C**. With the purpose of reducing the number of attributes to simplify the analysis of the dataset **C**, a subset $\mathbf{C_2}$ was formed by **C**. For this reason, the algorithm presented in [26] was used for attribute selection, it was necessary to designate a class label attribute, *motives for autopsy rejection* was selected for this objective. A summary of the analyzed datasets is presented in Table 3. The generated datasets do not have either missing, inconsistent or out of range values.

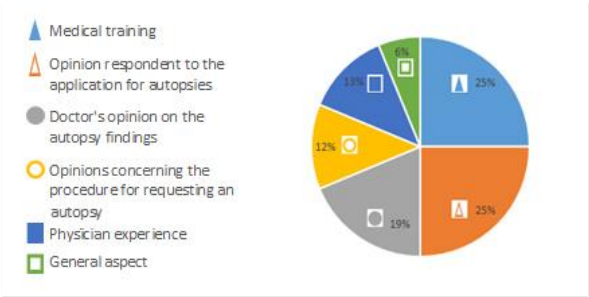


Fig. 1. Main areas explored by the survey.

Table 3. Data sets characteristics.

Set	Attributes	Objects
C	240	85
C_1	56	85
C_2	39	85

3.3 Application of the Data Mining Methods

Because of space limitations, only the association rules with greater confidence level obtained for each dataset are presented. Table 4 shows the association rules from the dataset **C**, the more frequent attributes in the rules are: *motives for autopsy acceptance*, *suitable staff to request an autopsy*, *motives for autopsy rejection by family* and *motives for not enough autopsies performed in the hospital*.

Table 4. Results of applying the Apriori algorithm to the dataset **C**.

Rule	Antecedent	Consequent	Confidence
1.	P06R01=S P09R04=S	P10R01=S	0.97
2.	P09R04=S P11R01=S	P10R01=S	0.96
3.	P11R01=S P14R01=S	P10R01=S	0.95
4.	P06R01=S P11R01=S	P10R01=S	0.94
5.	P09R04=S P10R01=S	P11R01=S	0.94
6.	P08R01=S	P10R01=S	0.93
7.	P11R05=S	P10R01=S	0.95
8.	P06R03=S	P10R01=S	0.95

The following analyzed dataset was **C_1**, the results are showed in Table 5 and they demonstrate the relations between attributes like: *motives for autopsy acceptance*, *suitable staff to request an autopsy*, and *efficient methods to request an autopsy*.

Table 5. Results of applying the Apriori algorithm to the dataset **C_1**.

Rule	Antecedent	Consequent	Confidence
9.	P06R01=S P12R03=S	P10R01=S	0.97
10.	P06R01=S P06R03=S	P10R01=S	0.94
11.	P06R01=S P11R01=S	P10R01=S	0.94
12.	P06R03=S P11R01=S	P10R01=S	0.94
13.	P06R01=S P11R05=S	P10R01=S	0.97

C_2 was analyzed with the objective of discovering interesting rules between the attributes with more correlations in the dataset and Table 6 shows the results. In this case the rules with more confidence were those that related *motives for autopsy rejection* and the *medical opinion about the discoveries in autopsies*.

Table 6. Results of applying the Apriori algorithm to the dataset C_2.

Rule	Antecedent	Consequent	Confidence
14.	P07R05=S	P10R01=S	1
15.	P04R02=S P07R03=S	P10R01=S	1
16.	P05R01=S	P10R01=S	1
17.	P07R03=S	P10R01=S	0.96
18.	P04R02=S P06R03=S P11R01=S	P11R01=S	0.95
19.	P01R05=S	P11R01=S	0.94
20.	P11R01=S P12R01=S	P10R01=S	0.96

The analysis of the data in general allowed the identification of the elements related to the autopsy according to the opinion of the medics of the hospital that was object of this study, as we can see in Table 7. The results are subject to a final interpretation of the specialist who designed the instrument used to collect the information, in order to confirm the precision of the generated models.

4 Conclusions

The prominent decrease in the number of autopsies in the hospitals around the world has raised questions about the motives for this phenomenon. The purpose of this work was analyze the possible causes of the reduction of autopsies in the hospital system of "Servicios de Salud de Veracruz" by means of association rule mining from the data that belongs to the medical opinions about such medical practice.

The analyzed data were collected through a survey that was applied to the doctors of the hospital. The survey focused on the medical opinions about the causes why the autopsies were not performed, the study level of the specialists, their years of experience, the cases of autopsies they have been involved on, among others.

Table 7. Preliminary results of the medical opinions about autopsies.

Aspects	Results
Motives for autopsy acceptance	-Establish definitive diagnoses -Wrong diagnoses -Pedagogical objectives -Interest
Efficient methods for request an autopsy	- Instrument that regulates the autopsies as obligatory for the patients who passed away in the hospital - Asking the consent to the family once the patient is admitted to the hospital - Request by the medic
Suitable staff to request an autopsy	- Medic
Motives for autopsy rejection	- Lack of interest - Refusal of the family - Known Disease
Discoveries in autopsies	- Originate in claim cases. - Originate in arbitration cases. - Cause discrepancy with the clinical diagnoses
Autopsy rejection by family	- Religious motives

The use of association rule mining techniques allowed perform a descriptive analysis of the problematic situation and find the correlations between the categorical attributes of the dataset, which formed the information obtained from the medical staff.

So, we can conclude that according to the survey respondents, the decrease in the number of autopsies in the analyzed hospital is because of the reduction in the requests, the rejection by family due to religious issues, death because of known diseases, and lack of interest for the family, as well as for the health staff in some cases.

Most of the medics identified the professional interest, the possibility of obtaining definitive diagnoses and the pedagogic usefulness as motivating elements to request an autopsy, procedure that must be initiated by the medics, this was demonstrated by several of the obtained rules. The rules emphasized the role of the medic in the request

of autopsies and identified as efficient methods for the request of an autopsy asking the consent of the family of the patient and the legalization by the institution, in a regional or state level, of the obligatory nature of the autopsy practice for the patients who die.

In this sense, as future work we suggest to study data of the clinical records of the patients who died in the hospital, to analyze with real data the trend of the causes that lead to perform autopsies in some patients and not in others. This will confirm the veracity of the results of this research. We also recommend perform similar studies in other parts of the country, to identify whether the medical opinions and the consequences of autopsies rejection differ by region. Finally, investigations that consider other data mining techniques, such as Bayesian Networks, can be seen as further work. This will establish through a comparative analysis which algorithms are more efficient for this kind of studies.

References

1. Sanz-Ortiz, J., Mayorga, M., Mart, A.: Autopsy in clinical oncology: Is it in crisis? Medicina Clínica, 137(7), pp. 317–320 (2011)
2. Oviedo Carrascal, E. A., Oviedo, A. I., Vélez Saldarriaga, G. L.: Minería de datos: aportes y tendencias en el servicio de salud de ciudades inteligentes. Revista Politécnica, 11(20), pp. 111–120 (2015)
3. Chao Wang, et al.: Exploration of the Association Rules Mining Technique for the Signal Detection of Adverse Drug Events in Spontaneous Reporting Systems. PLoS ONE, 7(7), pp. 6 (2012)
4. Timarán Pereira, R., et al.: La minería de datos aplicada al descubrimiento de patrones de supervivencia en mujeres con cáncer invasivo de cuello uterino. Revista Universidad y Salud, vol. 14, pp. 117–129 (2012)
5. Antonelli, D. et al., MeTA.: Characterization of Medical Treatments at Different Abstraction Levels. ACM Trans. Intell. Syst. Technol., no. 57, pp. 25 (2015)
6. Chih-Wen C., et al.: icuARM-II: improving the reliability of personalized risk prediction in pediatric intensive care units. In: Proceedings of the 5th ACM Conference on Bioinformatics, Computational Biology, and Health Informatics (BCB '14), ACM, New York, NY, USA, pp. 211–219 (2014)
7. Franco, A. M., Pérez, León Guzmán, E.:An approach to the risk analysis of diabetes mellitus type 2 in a health care provider entity of Colombia using business intelligence. In: Proc. of the 6th Euro American Conf. on Telematics and Information Systems (EATIS '12), Rogerio Patricio Chagas do Nascimento (Ed.). ACM, New York, NY, USA, pp. 49–56 (2012)
8. Shukor, N. A., Tasir, Z., Van der, H.: An Examination of Online Learning Effectiveness Using Data Mining. Procedia - Social and Behavioral Sciences, vol. 172, pp. 555–562 (2015)
9. Harwati, Permata, A., Ayu Wula, F.: Mapping Student's Performance Based on Data Mining Approach (A Case Study). Agriculture and Agricultural Science Procedia, vol. 3, pp. 173–177 (2015)
10. Yu, Z. J., Haghighat, F., Benjamin, C. M. F.: Advances and challenges in building engineering and data mining applications for energy-efficient communities. Sustainable Cities and Society (in press), pp. 2–6 (2015)

11. Gutierrez, B., et al.: A Bayesian network to predict coastal vulnerability to sea level rise. J. Geophys. Res. Earth Surf, vol. 116, pp. 15 (2011)
12. Bulteau, T., Baills, A., Petitjean, L., Garcin, M., Palanisamy, H.: Gaining insight into regional coastal changes on La Réunion island through a Bayesian data mining approach. Geomorphology, vol. 228, pp. 134–146 (2015)
13. Yousef, A. H.: Extracting software static defect models using data mining. Ain Shams Engineering Journal, vol. 6, pp. 133–144 (2015)
14. Adeniyi, D., Wei, Z., Yongquan, Y.: Automated web usage data mining and recommendation system using K-Nearest Neighbor (KNN) classification method. Applied Computing and Informatics, vol. 12, pp. 90–108 (2016)
15. Kaur, R., Singh, S.: A survey of data mining and social network analysis based anomaly detection techniques. Egyptian Informatics Journal (in press, corrected proof), no. 1 (2015)
16. Capozzoli, A., Grassi, D., Savino Pi, M.: Discovering Knowledge from a Residential Building Stock through Data Mining Analysis for Engineering Sustainability. Energy Procedia, vol. 83, pp. 370–379 (2015)
17. Tan, P.N., Steinbach, M., Kumar, V.: Introduction to Data Mining, Addison-Wesley, pp. 71–125 (2006)
18. Liu, B., Hsu, W., Ma, Y.: Integrating Classification and Association Rule Mining. In: Int. Conf. on Knowledge Disc. and Data Mining, pp. 80–86 (1998)
19. Bathla, H., et al.: Apriori Algorithm And Filtered Associattion Rule Mining. International Journal of Computer Science and Mobile Computing, vol. 4, pp. 209–306 (2015)
20. Han, J., Pei, J., Yin, Y.: Mining frequent patterns without candidate generation. In: Proceedings of the 2000 ACM-SIGMOD, pp. 1–12 (2000)
21. Srikant, R., Agrawal, R.: Mining Sequential Patterns: Generalizations and Performance Improvements. In: Proc. of the Int. Conf. on Extending Database Technology, pp. 1–17 (1996)
22. Scheffer, T.: Finding Association Rules that Trade Support Optimally against Confidence. In: 5th European Conference on Principles of Data Mining and Knowledge Discovery (2001)
23. Flach, P. A., Lachiche, N.: Confirmation-Guided Discovery of first-order rules with Tertius. Machine Learning, no. 42, pp. 61–95 (1999)
24. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between sets of items in large databases. In: Proceedings of the ACM SIGMOD Int. Conf. on Management of Data, Washington, DC, pp. 207–216 (1993)
25. Aher, S. B., Lobo, L.M.R.J.: A Comparative Study of Association Rule Algorithms for Course Recommender System in E-learning. International Journal of Computer Applications, vol. 39, no. 1, pp. 48–52 (2012)
26. Kira, K., Rendell, L. A.: A practical approach to feature selection. In: International Conference on Machine Learning, pp. 368–377 (1992)

ISSN 1870-4069

A Methodology for Optimizing the Process of Machining a Workpiece Using Multi-Objective Particle Swarm Optimization

Osslan Osiris Vergara Villegas, Carlos Felipe Ramírez Espinoza, Vianey Guadalupe Cruz Sánchez, Manuel Nandayapa, Jorge Luis García-Alcaraz

Universidad Autónoma de Ciudad Juárez, Instituto de Ingeniería y Tecnología, Ciudad Juarez, Chihuahua, México

{overgara, caramire, vianey.cruz, mnandaya, jorge.garcia}@uacj.mx

Abstract. In this paper a methodology based on multi-objective particle swarm optimization algorithm, for identifying the optimal parameters for machining a workpiece with a milling is presented. The time for machining, the material removal rate, and the feed rate were identified as the objectives to optimize. In addition, the proposal considered 4 constraints related to cutting tools, rotating speed of main spindle, cutting depth per passing, and speed interval for advance. Once the objective functions and constraints were determined, the test workpiece was designed by an unexperienced machinist by means of CATIA software, and then exported to Mastercam X in order to generate the G & M codes. The material selected for machining was delrin. In the experimentation stage, the methodology proposed was executed 50 times, and the parameters from the 2 best solutions were used to design 2 new workpieces. From the results obtained it was observed that the methodology proposed can support unexperienced operators in optimizing the parameters for machining. The machining time was reduced in 30%, material removal rate was increased about 55%, and an increment of 14% was obtained for the feed rate.

1 Introduction

Manufacturing is the process of converting, by hand or by machine, raw materials, components, or parts into finished goods that meet a customer's specifications. It implies, the application of physical and chemical processes to alter the geometry, properties, and appearance of a given starting material to make parts or products [1]. Manufacturing operations can be broadly classified into two different groups called processing, and assembly. The former, transforms a work material to the final desired product. The later, joins two or more components to create a new entity called an assembly or subassembly.

An important task inside processing operations is the material removal or well known as machining. Machining is the manufacturing process in which a cutting tool is used to remove the material excess from a workpiece, in such a way that the reminder obtains a desired shape. The machining process can be divided into two groups which

are: a) cutting process with traditional machining in which turning, milling, boring, and grinding are included, and b) cutting process with modern machining in which electrical discharge machining (EDM), and abrasive waterjet (AWJ) are included [2].

Typically, when a workpiece will be machined, a process planer selects the machining parameters based on his experience, and from the available handbooks. In the literature, a number of recommended machining parameters can be found, however, most of times those parameters do not suits exactly for a particular machine tool, material, and other combinations [3]. The poor selection of machining parameters leads to several drawbacks such as long machining times, and large waste of materials, therefore, it is imperative to enhance the parameters involved in workpiece machining. One of the main goals in machining is to produce high quality products with less cost and time constraints, and this can be due by optimizing the selection of process parameters such as cutting speed, depth of cut, feed rate, to mention a few. In several fields such as computer science, artificial intelligence, operation research, and manufacturing, optimization is the process of trying to find the best possible solution to a problem [4].

Several traditional and non-traditional methods can be used for determining the optimal machining parameters. Traditional techniques include dynamic programming, geometric programming, and deterministic techniques. However, these techniques either tend to result in local minima or take a long time to converge to an acceptable result [5]. On the other hand, non-traditional techniques include genetic algorithm (GA), simulated annealing (SA) [6], particle swarm optimization (PSO) [7], ant colony optimization (ACO) [8], and artificial bee colony (ABC) [9].

In the literature a number of single objective approaches have been presented for optimization of machining parameters such as material removal rate (MRR), surface roughness, cutting force, tool life, power consumption, to mention a few. However, several single objective problems must be applied to optimize the number of different criterions involved in machining processes. On the other hand, multi-objective approaches can be used to optimize several criterions at a time. In multi-objective optimization problems, the objectives to optimize are normally in conflict with respect to each other, therefore, there is no single solution to these problems. Hence, good trade-off solutions that represent the best possible compromises among objectives is frequently obtained. The trade-off solutions are called Pareto optimal set [10].

A reduced number of works, can be found in the literature regard multi-objective optimization of machining parameters. The paper of [11], considers the environmental dimension for optimizing milling cutting operations; three objectives, such as surface roughness, MRR, and cutting energy, were simultaneously optimized. A multi-objective optimization of lathe machining parameters for energy saving was shown in the work of [12], where three objectives including energy, cost, and quality were considered, which were affected by 3 constraints, namely cutting depth, feed rate, and cutting speed. A software prototype for solving multi-objective machining optimization problems was developed in [13]; the core was an algorithm based on exhaustive iterative search which guarantees the optimality of a determined solution in a given discrete search space, a wire electric discharge machining (WEDM), a micro WEDM, and a laser machining were considered for tests. The problem of computing the Pareto front was addressed in the paper of [14], by enumerating and characterizing 128 scenarios in sustainable machining operation involving 7 objectives including energy,

cost, time, power, cutting force, tool life and surface finish; several recommendations were offered to create a generic optimization framework. Finally, a method for complex optimization of cutting parameters with the objectives of energy efficiency and milling processing time, which integrates Taguchi method, response surface method (RSM), and multi-objective particle swarm optimization algorithm (MOPSO) was presented in [7].

Motivated from the revision above, and by observing that multi-objective methods have gained great attention, in this paper a method for multi-objective optimization of objective functions related to machining time, MRR, and feed rate for a milling machine is presented. The proposal considered 4 constraints related to rotating speed of main spindle, cutting depth per passing, speed interval for the advance, and cutting tools. Therefore, the primary focus of this paper is to suggest procedures to determine the optimal machining parameters for single-pass milling of a workpiece.

2 Proposed Methodology

The methodology proposed comprises two main stages; the first one, related to machining the workpiece, and the second one, related to MOPSO.

2.1 Machining the Workpiece

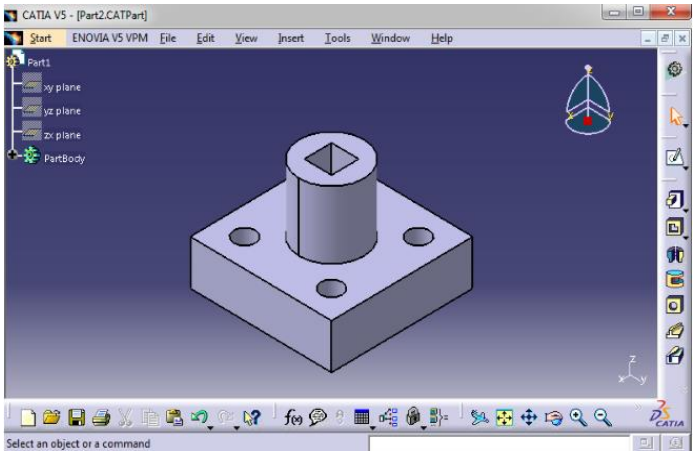
The first step in the methodology proposed, consists of machining a test workpiece with a specific geometry, by means of a computerized numerical control (CNC) machine. The milling selected was the vertical VIWA VF3KM400, which have a 5 horsepower (HP) triphasic motor, table dimensions of 1270 x 254mm, spindle career of 120mm, and variable speed range from 120-4200 revolutions per minute (rpm). A graphical example of the milling used is shown in Figure 1.

Fig. 1. The milling machine used for piece machining.

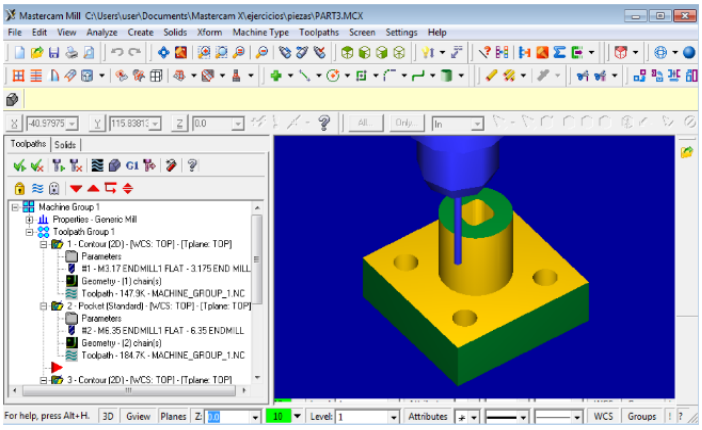
The material selected to build the workpiece was Polyoxymethylene (POM), better known as delrin, which was prepared by means of cutting, and carious operations. The cutting process started from a solid delrin block with a size of 65mm x 65mm x 50.8mm. The workpiece was made up of a cylinder placed at the center, with a diameter and a height of 30mm. The cylinder contains a hole of 12.5mm x 12.5mm with a depth

of 15mm. In addition, 4 holes were created in the corners of the piece, with a diameter and depth of 10mm.

The workpiece was designed by an unexperienced user by means of the computer aided three-dimensional interactive application (CATIA) software. The piece designed in CATIA was exported to Mastercam X in order to define the methodology, the tools, and the necessary parameters to perform the machining in a virtual way. The goal of this was to verify the existence of possible errors before the real workpiece machining. The workpiece designed in CATIA, and its correspondent representation in Mastercam X are shown in Figure 2.



a)



b)

Fig. 2. The workpiece designed in: a) CATIA, and b) Mastercam X.

Once that the correct design of the piece was virtually verified, the G and M codes were created, and used to control and drives the milling. The parameters used for the initial workpiece machining were set to: depth from 1.5 to 3mm, cutters with 2 and 4 teeth, with a size of 1/8 and 1/4 respectively, and 0.073 millimeters per tooth *mmPT*,

and rotating speed of main spindle of 3000 rpm. The final workpiece designed is shown in Figure 3.

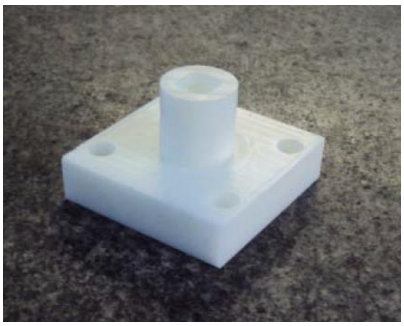


Fig. 3. The final workpiece designed.

2.2 Multi-Objective Particle Swarm Optimization

The metaheuristic used to solve the multi-objective optimization problem of this paper is known as PSO, that simulates the movements of a flock of birds which aim to find food. The main advantages of PSO includes better exploration and exploitation provided by local and global search capabilities of the algorithm.

According to [15], PSO starts with the random initialization of a population (swarm) of individuals (particles) in the *n*-dimensional search space (*n* is the dimension of problem). The particles fly over search space with adjusted velocities. Each particle keeps two values in its memory; its own best experience, and the best experience of the whole swarm.

2.2.1 Definition of the Objective Functions. Three different objective functions were defined in order to optimize the time for machining a piece (*TM*), the *MRR*, and the feed rate (*f*). It is important to mention that the three functions were merged in a sum function.

The function related to *TM* expressed in *min* is shown in equation 1:

$$TM = \sum_{n=1}^{T} t_n. \tag{1}$$

The function of *MRR* expressed in cm^3/min is shown in equation 2:

$$MR = n \times mmPT \times W \times A \times Z, \tag{2}$$

where *n* is the rotating speed of spindle measured in rpm, *mmPT* are the millimeters per tooth of the cutting tool, *W* is the radial width of cut, *A* is the axial depth of cut, and *Z* is the number of cutting tool teeth.

The third objective function corresponds to *f*, and it is expressed in *mm/min* as shown in equation 3:

$$f = n \times mmPT \times Z. \tag{3}$$

2.2.2 Constraints. The constraints defined for the proposed model were the following:

1. The rotating speed of main spindle must rank in an interval of 2500-4000 rpm, because it represents the middle to high rank of the milling machine.
2. The cutting depth per passing was defined in 1-4mm.
3. The speed interval for the advance was defined as 20-40 IPM.
4. Only high speed steel (HSS) cutting tools of 2 and 4 teeth were used, and the search space was limited to 1/16-5/16.

2.2.3 PSO Parameters. The parameters employed for the PSO algorithm were:

- The initial swarm was set to 10 particles.
- The maximum number of iterations was set to 50.
- The maximum and minimal speed [$-V_{max}$, V_{max}] interval was set to [-5, 5].
- The inertia factor decrement was realized with a maximum value of 1.4, and a minimal value of 0.4.
- The learning factors values of c_1 and c_2 were set to 1.49618.

3 Experimental Results

The initial parameters for machining the workpiece were inserted as inputs to PSO, and then the proposed algorithm was executed. Once that the stop criterion was met, an array containing the better solutions is delivered as an output.

For the experimentation, the two best solutions from the set obtained were selected, after that, by using the optimized parameters two new pieces were machined. The parameters obtained for the best solution (MOPSO1) were cutting depth of 3.5mm, cutting tool with a size of 1/4, 4 teeth, and 0.0656 *mmPT*, and rotating speed of 3388 rpm. For the second better solution (MOPSO2), the parameters obtained were cutting depth of 4mm, cutting tool with a size of 1/4, 4 teeth, and 0.0810 *mmPT*, and rotating speed of 3266 rpm. A summary of the values obtained from experimentation is shown in Table 1. In addition, the percentage improvements are shown in Table 2.

Table 1. Results obtained from experiments

Piece	*TM*	*MRR*	*f*
Original	18.03 *min*	12.554 *mm/min*	779.44 cm^3/min
MOPSO1	12.23 *min*	19.758 *mm/min*	889.01 cm^3/min
MOPSO2	12.47 *min*	18.709 *mm/min*	736.80 cm^3/min

Table 2. Improvement percentages from experiments

Piece	*TM*	*MRR*	*f*
Original	-	-	-
MOPSO1	32.06%	57.38%	14.06%
MOPSO2	30.84%	49.03%	-5.47%

3.1 Discussion

It should be noted from Tables 2 and 3 that the methodology proposed for machining a workpiece with optimized parameters offers competitive results. The TM was reduced in both solutions in 6 min (from 18 to 12), which represent a competitive percentage reduction of about 30%. The time reduction is mainly due to the increase of MRR, the faster the better for most machinist, therefore, the industry could deliver mores pieces by turn, and generate more economic gains.

The MRR was enhanced in 6 *mm/min*, which represent an approximate gain of 56%. From the experiments it was observed that even with the increase of MRR the cutting tool life is not affected (due to the material used), no chipping was presented, and a good surface finish is obtained. In addition, it is important to note that for the case of MOPSO1, *f* was enhanced from 779.44 cm^3/min to 889.01 cm^3/min, which represent a gain of 14.06%. However, as can be observed for the case of MOPSO2 no *f* improvement was obtained, unlike a decrement of 11 cm^3/min was obtained, even with this the other parameters influencing for good machining. In comparison with the literature works discussed in Section 1, the geometry of the workpiece machined in this paper is not trivial, therefore, the selection of the parameter by an unexperienced and even an experience machinist is complex. After the machining of the two workpieces with the parameters obtained with the methodology proposed, it was observed that its quality was good in terms of geometry, and surface finish.

4 Conclusions

In this paper a methodology based on multi-objective particle swarm optimization algorithm, for identifying the optimal parameters for machining a workpiece with a milling was presented. The results obtained from experiment proved that the optimized machining parameters of PSO could yield to improve the machining time, the MRR, and the feed rate for a particular geometry of a workpiece, and obtaining a good surface finish. The optimization results obtained in this paper confirm that the proposed optimization method is a very useful tool for multi-objective optimization of machining parameters. In addition, the proposed methodology can solve the trade-offs well when objectives were generally conflicting to each other, and constraints must be fulfilled.

Future work will be directed towards on extending the current approach to include more different and complex workpiece geometries. Also, it will be important extend the proposal to a wider of optimization of machining parameters. Regards to the material, it will be interesting perform tests with other material such as aluminum, cast iron, graphite stainless steel, among others. Finally, the model could be extended with objective functions related to cutting tool life, and surface roughness.

References

1. Groover, M.: Fundamentals of Modern Manufacturing: Materials, Processes and Systems. Wiley, New Jersey (2010)

2. Yusup, N., Mohd, A., Mohd, S.: Evolutionary Techniques in Optimizing Machining Parameters: Review and Recent Applications (2007–2011). Expert. Syst. Appl. 39, pp. 9909–9927 (2012)
3. Bharathi, S., Baskar, N.: Particle Swarm Optimization Technique for Determining Optimal Machining Parameters of Different Work Piece Materials in Turning Operation. Int. J. Adv. Manuf. Technol, 54, pp. 445–463 (2011)
4. Segura, C., Coello, C., Miranda, G., León, C.: Using Multi-Objective Evolutionary Algorithms for Single-objective Constrained and Unconstrained Optimization. Ann. Oper. Res., pp. 217–250 (2016)
5. Yildiz, A.: A New Hybrid Differential Evolution Algorithm for the Selection of Optimal Machining Parameters in Milling Operations. Appl. Soft. Comput., 13, pp. 1561–1566 (2013)
6. Mohd, A., Haron, H., Sharif, S.: Genetic Algorithm and Simulated Annealing to Estimate Optimal Process Parameters of the Abrasive Waterjet Machining. Eng. Comp., 27, pp. 251–259 (2011)
7. Li, C., Xiao, Q., Tang, Y., Li, L.: A method Integrating Taguchi, RSM and MOPSO to CNC Machining Parameters Optimization for Energy Saving. J. Clean. Prod., 135, pp. 263–275 (2016)
8. Kadirgama, K., Noor, M., Alla, A.: Response Ant Colony Optimization of End Milling Surface Roughness. Sensors, 10, pp. 2054–2063 (2010)
9. Samanta, S., Chakraborty, S.: Parametric Optimization of Some Non-Traditional Machining Processes using Artificial Bee Colony Algorithm. Eng. Appl. Art. Intell., 24, pp. 946–957 (2011)
10. Zavala, G., Nebro, A., Luna, F., Coello, C.: A survey of Multi-Objective Metaheuristics Applied to Structural Optimization. Struct. Multidisc. Optim., 49, pp. 537–558 (2014)
11. Yan, J., Li, L.: Multi-objective Optimization of Milling Parameters-the Trade-offs Between Energy, Production Rate and Cutting Quality. J. Clean. Prod., 52, pp. 462–471 (2013)
12. Wang, Q., Liu, F., Wang, X.: Multi-objective Optimization of Machining Parameters Considering Energy consumption. Int. J. Adv. Manuf. Technol., 71, pp. 1133–1142 (2014)
13. Kovacevic, M., Madic, M., Radovanovic, M., Rancic, D.: Software Prototype for Solving Multi-Objective Machining Optimization Problems: Application in Non-Conventional Machining Processes. Expert. Syst. Appl., 41, pp. 5657–5668 (2014)
14. Zhang, T., Owodunni, O., Gao, J.: Scenarios in Multi-objective Optimisation of Process Parameters for Sustainable Machining. Proceedia CIRP, 26, pp. 373–378 (2015)
15. Rezaee, A., Jasni, J.: Particle Swarm Optimization for Discrete Optimization Problems: A Review. Artif. Intell., 43, pp. 243–258 (2015)

ISSN 1870-4069

A Systemic Conceptual Model to Support Decision-Making in the Sustainability Assessment of Industrial Ecosystems

Dulce Rocío Mota-López, Cuauhtémoc Sánchez-Ramírez, Giner Alor-Hernández, Magno Ángel González Huerta

Division of Research and Postgraduate Studies, Instituto Tecnológico de Orizaba, Mexico

{csanchez, galor, magonzalez}@itorizaba.edu.mx, drociom@ito-depi.edu.mx

Abstract. One of the main challenges for companies is to operate in a sustainable way, that is, by protecting the ecosystems where they are located. As a result of such a challenge, different tools and models have been developed to assess their environmental impact and to help decision makers to define strategies to become more sustainable. However, most of these tools have concentrated on the internal parameters of the companies and have failed to assess the actual impact caused on the ecosystem. Therefore, this article proposes a conceptual model to analyze the environmental risks from companies and evaluate their impact on their ecosystems. Results obtained are: 1) a methodology to assess environmental risks and their impact on the environment and 2) a causal loop diagram that identifies the relationships between the industries and its environment, showing that the company's production rate is the key variable, since it determines the amount of both resources consumed and pollutants emitted, so its balance must be achieved in order to preserve the environment, which, otherwise, will collapse.

1 Introduction

Growth in population and economic activities – such as agriculture and industrialization – has negative effects on the environment. Therefore, preserving the balance in the exploitation and use of natural resources is crucial to avoid that such growth becomes unsustainable [1]. From this perspective, society, government institutions, and the industrial sector play crucial roles [2]. However, as regards industries, they are prominent actors because of three reasons [3]. First, as producers of goods and services, they are fundamental to satisfy the needs of both actual and future generations. Second, companies are vital to the economic development and wealth of regions. Third, their processes deteriorate the environment and contribute to the depletion of its natural resources.

Recent statistical data have revealed that the demand of natural resources such as water [4] and energy [5], from companies is increasing and will continue to increase. Such increment often exceeds the capacities of these resources to regenerate, not to mention that it also causes the emission of a greater amount of pollutants [6, 7].

Moreover, according to the World Water Assessment Programme [6], if companies maintain their current habits, by 2030, the world will face a 40% global water deficit.

This situation reveals that a robust approach must be adopted by industries in order to integrate sustainability into their processes and activities through concrete strategies and actions. Such an approach requires tools to assess the sustainability of these organizations [8] and to help decision and policy makers define the actions needed to reach a more sustainable status [9, 10].

According to the literature, the major progresses in industrial sustainable development have been achieved through eco-efficiency based tools [11, 12]. Unfortunately, these eco-efficiency based tools do not seem to provide positive results, since strategies employed merely evaluate parameters inside the company and do not consider the impact caused beyond the operational limits [3]. This implies that new approaches must be developed, and they must allow for a systemic analysis to measure the effectiveness of solutions proposed [13].

The main contribution of this research is therefore a model to assess the sustainability of companies from a systemic approach. The central part of such a model is the analysis of the interactions between the industrial sector and its environment based on water consumption, energy consumption emissions of air pollutants, and emissions of water pollutants. Such factors of environmental impact were selected due to the increasing consumption of water and energy and the high levels of hazardous substances emitted (section 3.1). Likewise, system dynamics (SD) was used to evaluate the model, since such a methodology enables to model different types of systems from the structural analysis of their variables, and this helps understand how systems work and the consequences that may arise from their status [14, 15].

As for the structure of this paper, the remaining sections are organized as follows. Section 2 provides an analysis of the most used tools and SD models for sustainability assessment in companies. Section 3 describes the proposed model in this research, while section 4 defines the methodology followed for its application. Eventually, section 5 introduces the causal loop diagram that depicts the interactions between the key variables from the conceptual model. Finally, section 6 presents conclusions and remarks for future work.

2 Tools for Industrial Sustainability Assessment

In the literature, a great number of industry-related sustainability assessment tools have been proposed and classified in various categories [2, 9, 16]. For example, [9] classified 48 tools in six categories: 1) individual/set of indicators, 2) composite indices, 3) socially responsible investment indices, 4) material and energy flow analysis, 5) life cycle analysis, and 6) environmental accounting. Likewise, [2] evaluates 16 tools, initiatives and methods for corporations to engage with sustainability. However, the analysis conducted to these tools shows that they do not appropriately reflect the damages caused to the ecological processes of ecosystems or the risks that such damages represent for the further development of companies, due to the fact that are focus on process or products. Tools such as SD, multi-criteria analysis, risk analysis, cost-benefit analysis and environmental impact assessment, allow supporting decision

making in sustainability assessment and have the potential to analyze the industry-ecosystem relations, although these tools are not only used for industrial purposes.

Thus, the characteristics that a tool must have to evaluate the environmental impact from companies are the following [3, 8, 11]:

1) **A holistic approach** to analyze as a whole all the elements comprised in the industrial and ecological systems.
2) **A large spatial scale** to adopt regional, national, and international scales without limiting to processes, plants, or products.

Table 1 shows the analysis of the mentioned characteristics in SD, multi-criteria analysis, risk analysis, cost-benefit analysis and environmental impact assessment. For instance, multi-criteria analysis, risk analysis and environmental impact assessment can employ large spatial scales, but they lack a holistic approach. From a different perspective, cost-benefit analysis does not rely on any of the three aforementioned characteristics, while SD includes them all, because it focuses on analyzing the relationships among the variables of the systems and allows for the use of different spatial scales [15, 17, 18]. Therefore, it is employed by this research.

Table 1. Characteristics of the sustainability assessment tools.

Authors	Tools	Characteristics	
		Holistic approach	Spatial scale
[17–19]	SD	X	X
[20–22]	Multi-criteria analysis	-	X
[23–25]	Risk analysis	-	X
[26, 27]	Cost-benefit analysis	-	-
[28–30]	Environmental impact assessment	-	X

2.1 SD Models for Sustainability Assessment

SD has been widely used to develop simulation models for sustainability assessment in different environmental, social, and economic systems [31].

In the literature several models holistically present the relationships that exist between environmental and socioeconomic systems, assessing the impact of domestic activities on water availability and quality [15, 32–36]; as well as on energy consumption and emissions of CO2 [37–41].

As for industrial simulation models, they have been applied to a wide range of industries, including the oil and gas sector [42], the automotive industry [43], the cement industry [44, 45], the steel industry [46], and the electric power industry [47]. Such models measure the use of natural resources and the emissions of hazardous substances, although they do not assess the impact caused on the environment.

Conceptual SD simulation models have also been proposed in the literature [14, 48, 49]. These models suggest a set of guidelines to construct them based on the three pillars of sustainability (economic, social and ecologic). In this case, the type of

environmental impact assessed depends on the user's criteria and the specific problem that is addressed.

Even though SD enables to model the different effects caused by social and economic activities on the environment, the literature review on simulation or conceptual SD models for sustainability assessment of companies shows that they do not integrate such damages caused to the environment.

3 Conceptual Model for Sustainability Assessment in Emerging Economies

This section analyzes the four environmental impact factors. These elements are part of the conceptual model proposed to evaluate sustainability, and they will be analyzed by means of a diagram that is part of the proposed model.

3.1 Environment and Industrial Impact Factors

Water consumption, energy consumption, emissions of water pollutants, and emissions of air pollutants, are important environmental impact factors[5, 50, 51] that contribute to the damage that the industrial sector causes to ecosystems. Their importance for industries is explained below.

Water consumption. Water is essential to industries because it is used for several purposes such as cleansing, heating, cooling, steam generation, ingredient, among others [6]. However, the rise in production and consumption -produced by population growth, urbanization and industrialization- has produced an increasing demand in freshwater, affecting it availability. Thus, it is a priority for industries to secure their access to the resource, especially in a competing environment. Besides, is expected that the world industrial water consumption will be increase by 400% in 2050.

Energy consumption. The industrial sector uses energy in activities such as product manufacturing, assembling, steam generation, heating, cooling, lighting of buildings among others; and is considered one of the main worldwide energy consumer [5]. According to the Organization for Economic Co-operation and Development (OECD), the amount of delivered energy in the industrial sector is expected to continue growing in the projection period 2010-2040 [52]. As regards the countries that are part of the OECD, the average increase per year is estimated to be 0.4%. However, for non-OECD members, the percentage increases up to 2.3% per year.

Emission of water pollutants. Industrial wastewaters are a serious ecological problem because they contain toxic substances that are hard to treat [4]. In developing countries around the 70% of the industrial wastewater produced is dumped untreated into water bodies, due to the lack of proper legal regulations and industrial investment in technology to treat them [53].

Emission of air pollutants. Industrial pollutants released in the air mainly involve carbon dioxide CO2. Such pollutants derive from the consumption of energy and the industrial processes of the companies [7]. Since the beginning of the industrial revolution, industrial CO2 emissions have been increasing, as a consequence of the surge in the use of fossil fuels. This tendency is expected to continue in the future despite the expansion in the use of alternative energy sources [54].

3.2 A Block Diagram of the Conceptual Model for Sustainability Assessment

Figure 1 shows the interactions between the company under assessment and the ecosystem, based in the environmental impact factors described above. As observed in Figure 1, these interactions work as follows:

1) The environment provides natural resources (water and energy) to the company for its production processes.
2) Companies produce goods and services according to the existing demand.
3) As sub-products of the production processes, air and water pollutants are generated and released into the environment.

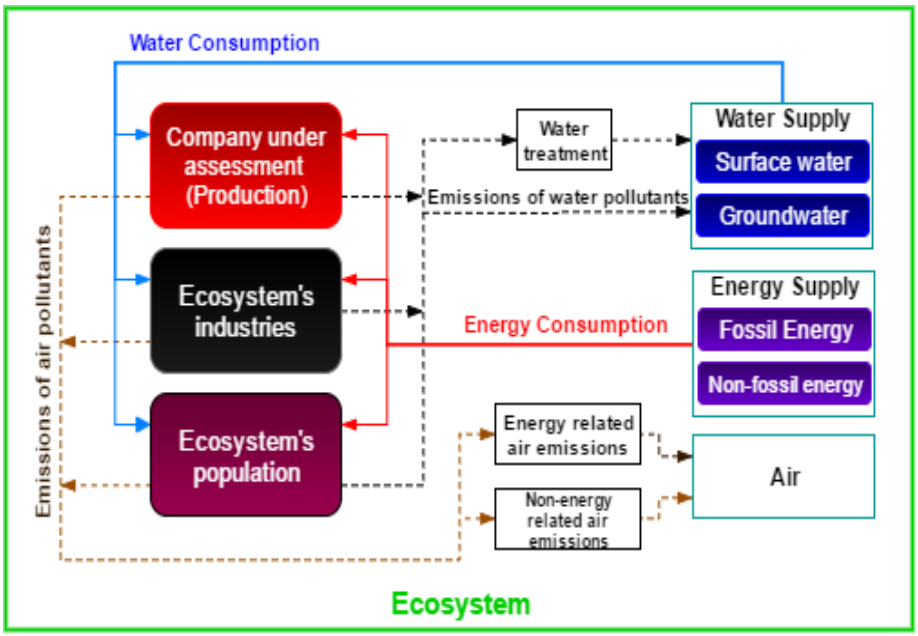


Fig. 1. Conceptual model.

The impact caused on the environment from the industries depends on the amount of natural resources consumed and on the amount of pollutants released and their levels of hazardousness. Thus, in order to improve the sustainability of a company a balance must exist to limit the amount of natural resources that the industry uses and the levels of pollutants that it emits. To guide industrial activities towards a more sustainable performance, one must define ecological limits for the use of natural resources and the emission of pollutants. Such limits ensure the correct functioning of the environment [9, 55], which is severely damaged when companies reach or are close to these limits. Moreover, if ecological limits are exceeded, they will seriously affect the ecological processes, which, in turn, will interrupt the industrial activities of companies involved [56]. Otherwise, companies will be responsible for severe water and energy shortages

and serious environmental degradation. However, a company is not the single actor that damages the environment, since activities from the population and other companies in the same ecosystem also have a negative effect on the environment. Therefore, both sectors are also integrated in the conceptual model.

Finally, it is worth mentioning that this model proposed to assess sustainability enables to know the current status of the environment by indicating the amount of damage caused by the company. From such a result, several strategies can be designed and proposed to diminish damages and guarantee the integrity of the environment and the future of where the model is applied.

4 Methodology Followed to Assess Sustainability

Figure 2 depicts the steps followed to assess the sustainability level of a company. On the one hand, the first stage is the characterization of the company to assess and its natural environment. On the other hand, the second step consists in developing the SD model. Each one of these stages is described below.

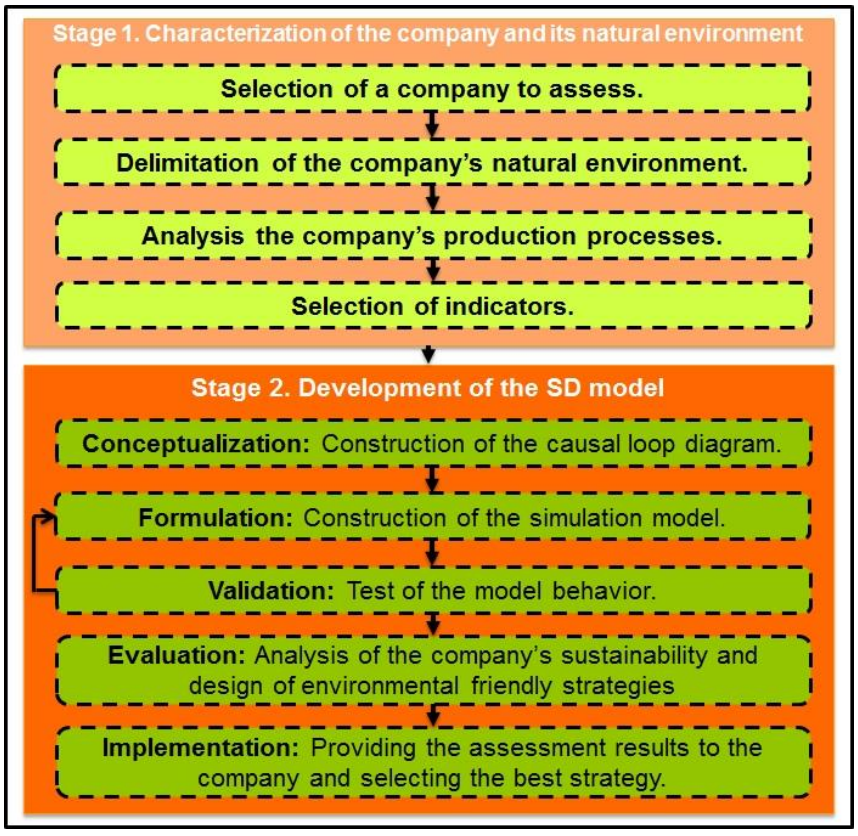


Fig. 2. Methodology to assess industrial ecosystem sustainability.

Stage 1: Characterization of the Company and its Natural Environment

This stage comprises four specific activities:

1. *Selection of a company to assess:* A company is selected to define the level of risk that it poses to the environment based on four factors: water consumption, energy consumption, emission of water pollutants, and emission of air pollutants.
2. *Delimitation of the company's natural environment:* Delimitating the natural environment affected by a company involves identifying: a) sources of supply of natural resources and processes involved in their availability, and b) sources of emission of air and water pollutants and their environmental impact.
3. *Analysis of the company's production processes:* This involves identifying the key processes of the company to determine the amount of both natural resources used and pollutants emitted.
4. *Selection of indicators:* Specific indicators will be selected to both assess the environmental impact of each one of the four factors (water consumption, energy consumption, emissions of water pollutants, and emissions of air pollutants) and define their ecological limits. Such limits will be used as reference to know the level of sustainability of the company.

Stage 2: Development of the SD Model

The model is developed following the stages described by Sterman (2000) and Cedillo-Campos and Sánchez-Ramírez (2008) for the creation of SD simulation models

1. *Conceptualization:* Key variables are selected. These variables represent the different interactions between a company and the environment. They are represented by a causal loop diagram.
2. *Formulation:* The simulation model (Forrester diagram) is developed based on the causal loop diagram. The system's behavior is described by means of mathematical equations related to the following factors: water consumption, energy consumption, emission of water pollutants, emission of air pollutants, generation of natural resources, and damages caused to these natural resources, among others,
3. *Validation:* The model is validated through different tests in order to determine whether its structure and behavior are consistent with the system that it represents [57, 58] (i.e.: the interactions between the industry and the environment). If results from these tests are not satisfactory, corrections must be made.
4. *Evaluation:* The level of sustainability of the company is assessed based on results obtained. Similarly, strategies for sustainability improvement are developed, and their effectiveness is tested. Finally, the most suitable strategy is selected.
5. *Implementation:* Final results are provided to the company and suggestions are offered to implement the most suitable strategy to improve sustainability.

This article presents the first step in the development of a SD model, which is the construction of the causal loop diagram.

5 Causal Loop Diagram

Table 2 lists and defines the variables that compose the causal loop diagram. Let us recall that such a diagram defines the different interactions between a company and the environment (Figure 3**¡Error! No se encuentra el origen de la referencia.**). Note that these variables were identified as key elements to assess the impact of the four environmental impact factors (water consumption, energy consumption, emissions of water pollutants, and emissions of air pollutants) over the environment.

Table 2. Variables of the causal loop diagram.

Authors	Variables	Description
[59, 60]	Demand	Amount of goods and services required by consumers.
	Company's production rate	Amount of products to manufacture.
[15, 34–36, 61, 62]	Company's water demand	Volume of water required for the industry's processes.
	Water consumption	Volume of water consumed in the region.
	Shortage	Environmental impact produced when the water consumption exceeds the available supply.
	Water supply	Total volume of water provided by surface and groundwater sources.
	Surface water	Volume of water withdrawn from superficial sources.
	Ground water	Volume of water withdrawn from groundwater sources.
	Company's emissions to water	Volume of wastewater produced by the company.
	Pollutant load	Degree of pollution of wastewater.
	Water quality	Measure of the water's physical and chemical properties that indicates the environmental impact caused to the resource.
	Company's water quality requirements	Parameter that defines the water quality level required for the company's processes.
	Quality satisfaction	Measure that indicates if the ecosystem's water quality complies with the industry's water quality requirements.
	Water with the quality requirements	Volume of water that complies with the quality requirements.
	Water treatment	Volume of water treated to remove pollutants.
	Water availability for the company	Volume of water available for the company's processes.
[46, 63]	Energy consumption	Amount of energy used for the company's processes.
	Energy supply	Amount of energy supplied to the industry.
	Energy security	Measure that indicates if the company's energy supply is reliable or not.
[39, 44, 45, 64]	Fossil fuel energy use	Amount of energy from fossil fuels.
	Emissions to air related to fossil fuel energy use	Emissions of air pollutants produced by the use of fossil fuels.
	Company's emissions to air	Pollutant substances released by the company to the atmosphere.
	Air quality	Measure of the amount of pollutants in the atmosphere that indicates the environmental impact produced to the air.
	Pressure to reduce emissions	Actions taken to promote the reduction of pollutant emissions to air.
	Adoption of environmental friendly alternatives	Alternatives that lead to the reduction of emissions of air pollutants.
[36, 61]	Ecosystem's industries water demand	Volume of water required by the industries and the population in the ecosystem.
	Ecosystem's population water demand	
	Ecosystem's industries emissions to water	Volume of wastewater produced by the industries and the population in the ecosystem.
	Ecosystem's population emissions to water	
	Ecosystem's industries energy consumption	Amount of energy required by the industries and the population in the ecosystem.
	Ecosystem's population energy consumption	
	Ecosystem's industries emissions to air	Volume of air pollutants emissions produced by the industries and the population in the ecosystem.
	Ecosystem's population air emissions	

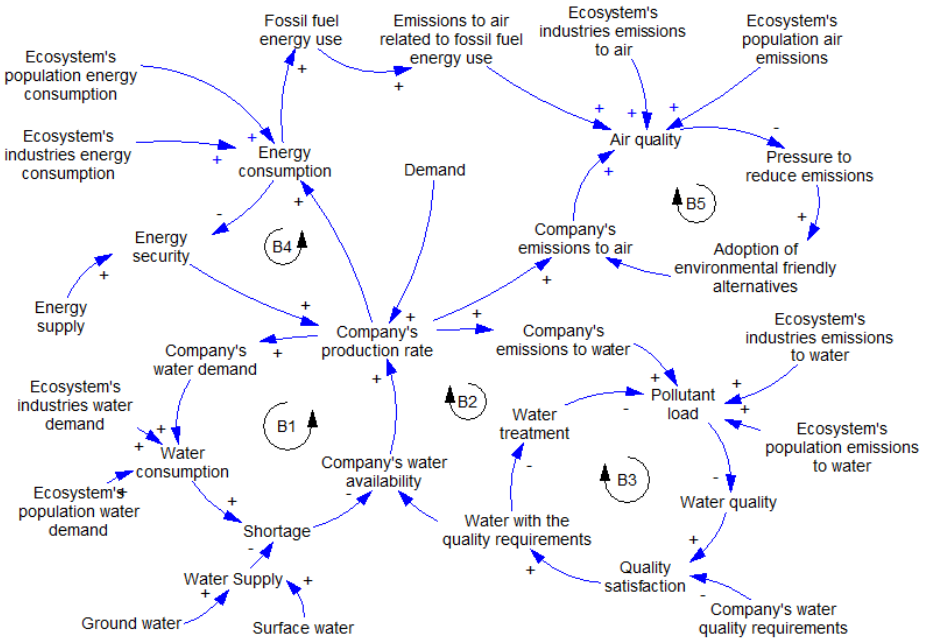


Fig. 3. Causal loop diagram.

The feedback loops from the causal loop diagram that show the interactions produced by the environmental impact factors are analyzed are the following:

B1: The company's production rate defines to a great extent the amount of water that every company consumes. When this factor causes its excessive use, the total water consumption in the environment increases. Therefore, in this case, and within a scenario of industrial growth, water becomes insufficient for the company, since its consumption exceeds its availability.

B5: The company's production rate and the industry sector define to the greatest extent the type and volume of industrial wastewater discharged by companies, since it determines the toxicity and concentration of pollutants found in the discharges. Moreover, the quality of water is essential to the industrial sector, because all industrial processes require specific quality levels of such a resource. Therefore, if these levels are too high, sources to obtain water become scarce, and water available is insufficient.

B7: Energy security allows companies to ensure the correct operation of their production processes at all stages and successfully meet certain requirements. However, the energy security in the industrial sector depends on how much it is consumed. Thus, if the company increases its use of electricity due to higher production rates, energy security will be compromised, and its availability may be interrupted.

B8: Air quality diminishes when the company emits high levels of air pollutants. Such chemicals are the result of several industrial processes and the use of fossil fuels. Therefore, since poor air quality negatively affects human health and ecosystems, the company must develop strategies to reduce their emissions.

The global analysis of the causal loop diagram shows that the company's production rate is the key variable that influences the four environmental impact factors, and it also becomes affected by those factors. This means that if at least one of them is out of control, the company and their processes will be seriously affected, and their future will be compromised. Therefore, this research highlights that it is important to strive to maintain and ensure a balance between the consumption of natural resources and the emission of pollutants.

6 Conclusions and Future Work

By using the SD approach and a causal loop diagram, this research proposes a model that analyzes the following aspects: 1) the uncontrolled interactions between a company and the ecosystem, which are produced by four environmental impact factors: water consumption, energy consumption, emission of water pollutants, and emission of air pollutants, and 2) SD is a tool that enable to model in an integrated manner the different interactions between the company and the ecosystem, because it follows an holistic approach, which allows for the analysis of larger spatial and temporal scales.

Likewise, this new model provides the foundations to construct simulation models for sustainability assessment that efficiently provide information to develop strategies that ensure the conservation of the environment. Similarly, the model demonstrates that environmental damages contribute to a difficult regeneration of natural resources, which hence limits the development and correct operation of companies.

As future work, the conceptual model proposed will be applied in a case study by following the stages described in the methodology section, with the aim of validating the model in its field of application and obtaining feedback about it. Finally, the usefulness of this model can be improved by integrating more environmental impact factors, such as the soil use and the generation of solid waste, among others. This will provide a better understanding and a more complete view of the damages that companies cause to the ecosystems.

References

1. Hay, L., Duffy, A., Whitfield, R.I.: The Sustainability Cycle and Loop: Models for a more unified understanding of sustainability. J. Environ. Manage, 133, pp. 232–257 (2014)
2. Lozano, R.: Towards better embedding sustainability into companies' systems: An analysis of voluntary corporate initiatives. J. Clean. Prod., 25, pp. 14–26 (2012)
3. Kunz, N.C., Moran, C.J., Kastelle, T.: Conceptualising "coupling" for sustainability implementation in the industrial sector: A review of the field and projection of future research opportunities (2013)
4. WWAP: The United Nations World Water Development Report 4: Managing Water under Uncertainty and Risk, Vol.1, Knowledge Base, Vol. 2, Facing Challenges, Vol. 3 (2012)
5. IEA: International Energy Outlook 2014. , Washington, D.C. (2014)
6. WWAP: The United Nations World Water Development Report: Water for a Sustainable World, Paris (2015)
7. IEA: CO2 Emissions From Fuel Combustion Highlights, https://www.iea.org/publications/freepublications/publication/CO2EmissionsFromFuelCombustionHighlights2014.pdf

8. Ness, B., Urbel-Piirsalu, E., Anderberg, S., Olsson, L.: Categorising tools for sustainability assessment. Ecol. Econ, 60, pp. 498–508 (2007)
9. Angelakoglou, K., Gaidajis, G.: A review of methods contributing to the assessment of the environmental sustainability of industrial systems. J. Clean. Prod (2015)
10. Singh, R.K., Murty, H.R., Gupta, S.K., Dikshit, A.K.: An overview of sustainability assessment methodologies. Ecol. Indic., 9, pp. 189–212 (2009)
11. Bocken, N.M.P., Short, S.W., Rana, P., Evans, S.: A literature and practice review to identify Sustainable Business Model Element Archetypes. J. Clean. Prod., 65, pp. 42–56 (2014)
12. Huysman, S., Sala, S., Mancini, L., Ardente, F., Alvarenga, R.A.F., De Meester, S., Mathieux, F., Dewulf, J.: Toward a systematized framework for resource efficiency indicators. Resour. Conserv. Recycl., 95, pp. 68–76 (2015)
13. Hjorth, P., Bagheri, A.: Navigating towards sustainable development: A system dynamics approach. Futures, 38, pp. 74–92 (2006)
14. Musango, J.K., Brent, A.C.: A conceptual framework for energy technology sustainability assessment. Energy Sustain. Dev., 15, pp. 84–91 (2011)
15. Zhang, Z., Lu, W.X., Zhao, Y., Song, W.B.: Development tendency analysis and evaluation of the water ecological carrying capacity in the Siping area of Jilin Province in China based on system dynamics and analytic hierarchy process. Ecol. Modell, 275, pp. 9–21 (2014)
16. Moldavska, A., Welo, T.: On the Applicability of Sustainability Assessment Tools in Manufacturing. Procedia CIRP, 29, pp. 621–626 (2015)
17. Forrester, J.W.: Industrial Dynamics. Harv. Bus., Rev. 36, pp. 37–66 (1958)
18. Sterman, J.D.: Business dynamics: Systems thinking and modeling for a complex world. Irwin McGraw-Hill (2000)
19. Cedillo Campos, M.G., Sánchez Ramírez, C.: Análisis Dinámico de Sistemas Industriales. Trillas (2008)
20. Amrina, E., Ramadhani, C., Vilsi, A.L.: A Fuzzy Multi Criteria Approach for Sustainable Manufacturing Evaluation in Cement Industry. Procedia CIRP, 40, pp. 620–625 (2016)
21. Gupta, S., Dangayach, G.S., Singh, A.K., Rao, P.N.: Analytic Hierarchy Process (AHP) Model for Evaluating Sustainable Manufacturing Practices in Indian Electrical Panel Industries. Procedia - Soc. Behav. Sci., 189, pp. 208–216 (2015)
22. Herva, M., Roca, E.: Review of combined approaches and multi-criteria analysis for corporate environmental evaluation. J. Clean. Prod., 39, pp. 355–371 (2013)
23. Li, D., Zhang, C., Pizzol, L., Critto, A., Zhang, H., Lv, S., Marcomini, A.: Regional risk assessment approaches to land planning for industrial polluted areas in China: The Hulunbeier region case study. Environ. Int., 65, pp. 16–32 (2014)
24. Goerlandt, F., Reniers, G.: On the assessment of uncertainty in risk diagrams. Saf. Sci., 84, pp. 67–77 (2016)
25. Izvercian, M., Ivascu, L.: Semantic Assessment of the Risk: A Multidimensional Data Model for Sustainable Enterprise. Procedia Technol., 16, pp. 598–605 (2014)
26. Finnie, B., Stuart, J., Gibson, L., Zabriskie, F.: Balancing environmental and industry sustainability: A case study of the US gold mining industry. J. Environ. Manage, 90, pp. 3690–3699 (2009)
27. Araújo, C., Almeida, M., Bragança, L., Barbosa, J.A.: Cost–benefit analysis method for building solutions. Appl. Energy, 173, pp. 124–133 (2016)
28. Kobayashi, Y., Peters, G.M., Khan, S.J.: Towards More Holistic Environmental Impact Assessment: Hybridisation of Life Cycle Assessment and Quantitative Risk Assessment. Procedia CIRP, 29, pp. 378–383 (2015)
29. Naddeo, V., Belgiorno, V., Zarra, T., Scannapieco, D.: Dynamic and embedded evaluation procedure for strategic environmental assessment. Land use policy, 31, pp. 605–612 (2013)
30. Kusi-Sarpong, S., Sarkis, J., Wang, X.: Assessing green supply chain practices in the Ghanaian mining industry: A framework and evaluation. Int. J. Prod. Econ., pp. 1–17

(2015)
31. Forrester, J.W.: System dynamics—a personal view of the first fifty years. Syst. Dyn., Rev. 23, pp. 345–358 (2007)
32. Chung, G., Lansey, K., Blowers, P., Brooks, P., Ela, W., Stewart, S., Wilson, P.: A general water supply planning model: Evaluation of decentralized treatment. Environ. Model. Softw., 23, pp. 893–905 (2008)
33. Rehan, R., Knight, M. a, Unger, A. J. A., Haas, C.T.: Financially sustainable management strategies for urban wastewater collection infrastructure – development of a system dynamics model, 7, (2013)
34. Zarghami, M., Akbariyeh, S.: System dynamics modeling for complex urban water systems: Application to the city of Tabriz, Iran. Resour. Conserv. Recycl., 60, pp. 99–106 (2012)
35. Sušnik, J., Vamvakeridou-Lyroudia, L.S., Savić, D.A., Kapelan, Z.: Integrated System Dynamics Modelling for water scarcity assessment: Case study of the Kairouan region. Sci. Total Environ., 440, pp. 290–306 (2012)
36. Mavrommati, G., Bithas, K., Panayiotidis, P.: Operationalizing sustainability in urban coastal systems: A system dynamics analysis. Water Res., 47, pp. 7235–7250 (2013)
37. Guan, D., Gao, W., Su, W., Li, H., Hokao, K.: Modeling and dynamic assessment of urban economy-resource-environment system with a coupled system dynamics - Geographic information system model. Ecol. Indic., 11, pp. 1333–1344 (2011)
38. Jin, W., Xu, L., Yang, Z.: Modeling a policy making framework for urban sustainability: Incorporating system dynamics into the Ecological Footprint. Ecol. Econ., 68, pp. 2938–2949 (2009)
39. Feng, Y.Y., Chen, S.Q., Zhang, L.X.: System dynamics modeling for urban energy consumption and CO2 emissions: A case study of Beijing, China. Ecol. Modell., 252, pp. 44–52 (2013)
40. Wu, Z., Xu, J.: Predicting and optimization of energy consumption using system dynamics-fuzzy multiple objective programming in world heritage areas. Energy, 49, pp. 19–31 (2013)
41. Robalino-López, A., Mena-Nieto, A., García-Ramos, J.E.: System dynamics modeling for renewable energy and CO2 emissions: A case study of Ecuador. Energy Sustain. Dev., 20, pp. 11–20 (2014)
42. Kiani, B., Ali Pourfakhraei, M.: A system dynamic model for production and consumption policy in Iran oil and gas sector. Energy Policy, 38, pp. 7764–7774 (2010)
43. Abduaziz, O., Cheng, J.K., Tahar, R.M., Varma, R.: A hybrid Simulation model for Green Logistics Assessment in Automotive Industry. Procedia Eng., 100, pp. 960–969 (2015)
44. Ansari, N., Seifi, A.: A system dynamics model for analyzing energy consumption and CO2 emission in Iranian cement industry under various production and export scenarios. Energy Policy, 58, pp. 75–89 (2013)
45. Anand, S., Vrat, P., Dahiya, R.P.: Application of a system dynamics approach for assessment and mitigation of CO2 emissions from the cement industry. J. Environ. Manage, 79, pp. 383–398 (2006)
46. Ansari, N., Seifi, A.: A system dynamics analysis of energy consumption and corrective policies in Iranian iron and steel industry. Energy, 43, pp. 334–343 (2012)
47. Saysel, A.K., Hekimoğlu, M.: Exploring the options for carbon dioxide mitigation in Turkish electric power industry: System dynamics approach. Energy Policy, 60, pp. 675–686 (2013)
48. Kibira, D., Jain, S., McLean, C.R.: A System Dynamics Modeling Framework for Sustainable Manufacturing (2009)
49. Zhang, H., Calvo-Amodio, J., Haapala, K.R.: A conceptual model for assisting sustainable manufacturing through system dynamics. J. Manuf. Syst., 32, pp. 543–549 (2013)
50. WWPA: The United Nations World Water Development Report 2015: Water for a

Sustainable World, Paris (2012).
51. IEA: Worldwide Trends in Energy Efficiency: Key Insights from IEA Indicator Analysis, https://www.iea.org/publications/freepublications/publication/Indicators_2008.pdf
52. IEA: CO2 Emissions From Fuel Combustion Highlights, https://www.iea.org/publications/freepublications/publication/CO2EmissionsFromFuelCombustionHighlights2014.pdf
53. UN-Water: Wastewater Management – A UN-Water analytical Brief, http://www.unwater.org/publications/publications-detail/en/c/275896
54. IEA: Tracking Industrial Energy Efficiency and CO2 Emissions, https://www.iea.org/publications/freepublications/publication/tracking_emissions.pdf
55. Bithas, K.: Tracing operational conditions for the Ecologically Sustainable Economic Development: The Pareto optimality and the preservation of the biological crucial levels. Environ. Dev. Sustain, 10, pp. 373–390 (2008)
56. Bithas, K., Nijkamp, P.: Biological integrity as a prerequisite for sustainable development : A bioeconomic perspective Research Memorandum 2013-9 Kostas Bithas (2013)
57. Forrester, J.W., Senge, P.M.: Tests for building confidence in system dynamics models, (1980)
58. Qudrat-Ullah, H., Seong, B.S.: How to do structural validity of a system dynamics type simulation model: The case of an energy policy model. Energy Policy, 38, pp. 2216–2224 (2010)
59. Musango, J.K., Brent, A.C., Amigun, B., Pretorius, L., Müller, H.: A system dynamics approach to technology sustainability assessment: The case of biodiesel developments in South Africa. Technovation, 32, pp. 639–651 (2012)
60. Al-Sharrah, G., Elkamel, A., Almanssoor, A.: Sustainability indicators for decision-making and optimisation in the process industry: The case of the petrochemical industry. Chem. Eng. Sci., 65, pp. 1452–1461 (2010)
61. Liu, H., Benoit, G., Liu, T., Liu, Y., Guo, H.: An integrated system dynamics model developed for managing lake water quality at the watershed scale. J. Environ. Manage, 155, pp. 11–23 (2015)
62. Madani, K., Mariño, M.A.: System dynamics analysis for managing Iran's Zayandeh-rud river basin. Water Resour. Manag., 23, pp. 2163–2187 (2009)
63. Aslani, A., Helo, P., Naaranoja, M.: Role of renewable energy policies in energy dependency in Finland: System dynamics approach. Appl. Energy, 113, pp. 758–765 (2014)
64. Lee, S., Geum, Y., Lee, H., Park, Y.: Dynamic and multidimensional measurement of product-service system (PSS) sustainability: A triple bottom line (TBL)-based system dynamics approach. J. Clean. Prod., 32, pp. 173–182 (2012)

ISSN 1870-4069

Coverage Analysis of Fire Stations Located in Puebla City and Determination of New Possible Required Locations

Lucero Alejandra Salazar Jáuregui, Diana Sánchez-Partida, José Luis Martinez-Flores

Popular Autonomous University of the State of Puebla (UPAEP University), Puebla, Mexico

{Luceroalejandra.salazar, Diana.sanchez, Joseluis.martinez01}@upaep.mx

Abstract. The growing population in Puebla City these last couple of years entailed the expansion of its original urban area. The most recent settlements have established on areas that are highly exposed to natural disasters, thus, requiring the proper emergency service coverage provided by Fire Stations. To assure this kind of service coverage, the current fire stations in Puebla City were evaluated through mapping tools and given parameters, in order to recognize the capacity to satisfy the entire city's service demand. Based on the coverage analysis it was determined the need for an extra facility. The characteristics of the problem suggested that the most appropriated model to use is the Maximum Coverage Model, which is adapted for Lingo software to determine the optimal location to place the new fire station, looking for the full or largest possible coverage.

1 Introduction

The expansion of the urban area of Puebla City began in the middle of the twentieth century as a result of the exponential population growth. According to the National Institute of Statistics and Geography (INEGI), in the last 10 years the population in Puebla has increased by 12% and it shows a trend to settle up towards the southern part of the city. Due to topographic, climatologic and urban design characteristics, many of these new settlements are highly vulnerable to different types of risks caused mainly by natural disasters such as mudslides, fires and floods. The Fire Department of Puebla City has implied that the vulnerability for these new urban areas could be higher, considering that the response time for emergencies might take longer due to the lack of closer fire stations.

As stated by the Public Security Law of Puebla State, in case of an emergency, the first unit to provide the service must arrive within the first 10 minutes after a call. Although, if the service required is not a priority or there are not available resources to cover the service immediately, the response time could take between 15 and 25 minutes. Nowadays, the emergencies that occur in these new communities are covered by any of the current seven fire stations located in the main urban area; however, due to the distances between these communities and the nearest fire stations, the emergency response time could take up to 50 minutes.

The Fire Department of Puebla recognizes the need for more fire stations that can provide the emergency service for these communities within a fewer response time. Nonetheless, the low population density of these communities (due to its recent creation) along with the limited government budget, would not justify installing more than one fire station to provide the required service.

The objective of this research is to analyze the coverage scope of the existing fire stations in Puebla City to be able to identify the service uncovered areas and to determine the optimal location for a new facility, ensuring the full or maximum possible service coverage for the areas that require it.

2 Literature Review

At present, mapping tools have become a strong support for facility location decisions, especially with the integration of Geographic information systems (GIS).

According to IFMA Foundation [1] "A GIS can be used by facility managers for space management, visualization and planning, and emergency and disaster planning and response, as well as many other applications"

Resent research related to public facilities location relies on the use of two of the main mapping applications that integrate GIS: Esri and Google Maps.

Mosquera [2], in her study of *Optimal areas determination for new fire stations located at Bogota based on special analysis tools*, combines the use of ESRI with the Analytic Hierarchy Process (AHP), while other authors such as Gu et al [3] employs the use of Google Maps to estimate travelling distances and times in their study *Optimization of preventive health care facility locations.*

Since Google Maps is a widely known and a user friendly app, we use this web-mapping tool to generate part of the data that will be the input for the facility location model. To determine the mathematical model to be used, we refer and briefly describe the most commonly used models for emergency facility locations:

- P-Median model: The objective function is to minimize the average distance between facilities and demanding nodes.
- P-Center model: The objective function is to minimize the maximum distance/time between a facility and the demanding nodes, placing P number of facilities required.
- Set Covering Model: The objective function is to minimize the cost of placing the required facilities while ensuring that all demanding nodes are satisfied.

Based on the problem features, the model to determine the optimal location for the new fire station must have in consideration the budget restriction. A variation of the Set covering model that fits this requirement is the Maximum Coverage Model.

As quoted by Church and ReVelle [4], "The maximal covering location problem was formulated to address planning situations which have an upper limit on the number of facilities to be sited. The objective of the MCLP is to locate a predetermined number of facilities, p, in such a way as to maximize the demand that is covered".

For this case study, we will use this last model; however, we will not take into account values of demand, as a demanding point will be considered as served as long as it is within a coverage range of any fire station.

3 Methodology

3.1 Study Area

The City of Puebla has currently seven fire stations mainly distributed on the central urban area. Figure 1 shows the locations of these fire stations, along with the southern area tentatively considered as uncovered (marked in pink). It also shows two recognized sets of communities in this area. Set 1 includes 5 communities located to the north of a river (Balsas River) and Set 2 includes 5 communities located to the south of the river.

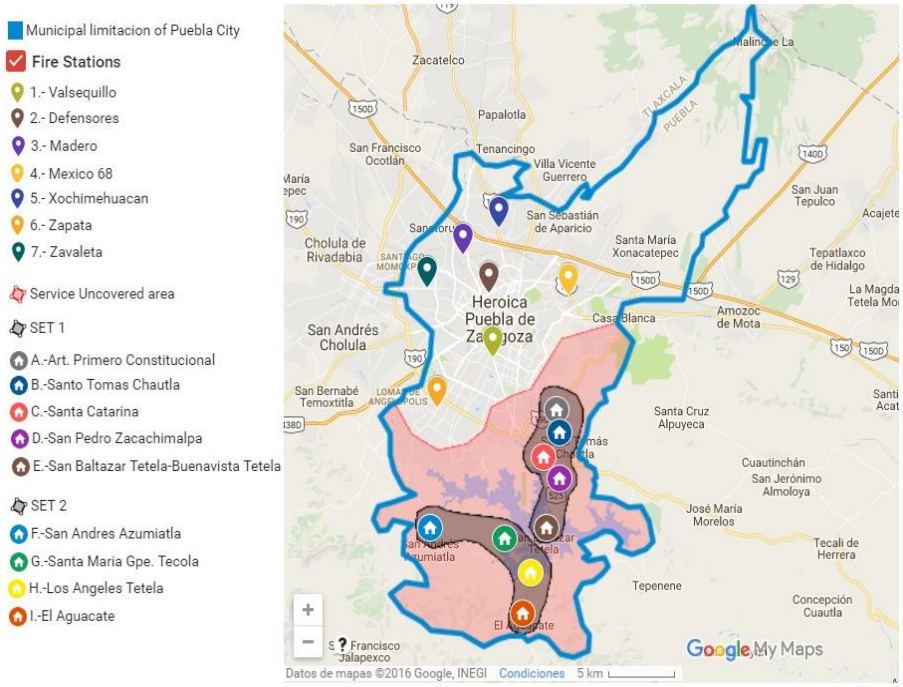


Figure 1. Fire stations, uncovered area and communities within Puebla City.

The coverage analysis consists in detecting the capacity of the current fire stations to satisfy the demand of emergency service for the entire city, considering strictly the law restriction of 10 minutes-first unit arrival. To do so, we need to define firstly the measurement units we will use for the data. Google Maps provides accurate distances between points, as well as estimated times. The estimated times are relative to speed limits and current traffic among others conditions, despite this, fire trucks are allowed to skip some of these laws for emergency situations. Therefore, the measurement unit for data will be only distances. It will be assumed that fire trucks travel at least 1 km per minute (considering circumstances that may slow it down), so the coverage limit per fire station will be 10 kilometers (based on the 10 minutes-first unit arrival restriction). The analysis will be developed in two parts. First it will be the analyzed the main central urbanized area and then, the south possible uncovered area.

3.2 Coverage Analysis

To analyze the coverage in the central urbanized area, we suppose each fire station represents a node in a network and the arcs of the network are the distances between each fire station. As each node should cover a maximum distance of 10 km, we must assure that the arcs between each node and the closest one is at least 20 km (10 km per node). The matrix of distances between each node is shown in Table 1.

Table 1. Distances between fire stations located in Puebla City.

Fire Station From-to matrix	Valsequillo	Defensores	Madero	Mexico 68	Xochime-huacan	Zapata	Zavaleta
Valsequillo		17km	22km	23km	26km	18km	23km
Defensores	17km		11km	14km	13km	25km	18km
Madero	22km	11km		17km	11km	23km	13km
Mexico 68	23km	14km	17km		16km	33km	31km
Xochimehuacan	26km	13km	11km	16km		31km	21km
Zapata	18km	25km	23km	33km	31km		15km
Zavaleta	23km	18km	13km	31km	21km	15km	

Based on the matrix data, it is possible to recognize that each fire station has at least another fire station located within the limit of 20 km, therefore it can be assumed that the current set of fire stations meets the conditions of response coverage for this area. Unlike the central area, the southern area distances matrix (shown in table 2) only considers the distances between each of the communities and the two nearest fire stations (Zapata and Valsequillo).

Table 2. Distances between Zapata and Valsequillo fire stations and the communities in the south area.

Community / Fire Station	Zapata	Valsequillo
San Andres Azumiatla	11km	18km
Santa Maria Gpe. Tecola	17km	23km
Los Angeles Tetela	21km	28km
El Aguacate	27km	35km
San Baltazar Tetela	17km	15km
San Pedro Zacachimalpa	15km	12km
Santa Catarina	11km	10km
Santo Tomas Chautla	12km	9km
Art. Primero Constitucional	13km	9km

As seen on the estimations in table 2, some of the 9 communities can be considered as covered because they are located within the 10 kilometers condition to any of the

two nearest fire stations. However, none of these fire stations seems to be able to provide the service within the 10 minutes restriction for all the communities located in this southern area. Therefore, a facility will have to be located closer to these communities to provide the emergency service. It is important to mention that due to the need of utilities such as electricity, gas and water, among others, the facility must be located within a candidate community and not in an average geographical point. In this way, we proceed to generate the distance matrix for all the communities located in this area:

Table 3. Distances Matrix between each of the communities in the south area of the municipal.

From-to matrix	San Andres Azumiatla	Santa Maria Gpe. Tecola	Los Angeles Tetela	El Aguacate	San Baltazar Tetela	San Pedro Zacachimalpa	Santa Catarina	Santo Tomas Chautla	Art. Primero Constitucional
San Andres Azumiatla		7km	12km	18km	27km	25km	22km	22km	24km
Santa Maria Gpe. Tecola	7km		16km	10km	16km	34km	27km	28km	29km
Los Angeles Tetela	12km	16km		7km	7km	25km	26km	26km	30km
El Aguacate	18km	10km	7km		8km	26km	27km	28km	31km
San Baltazar Tetela	27km	16km	7km	8km		4km	8km	8km	12km
San Pedro Zacachimalpa	25km	34km	25km	26km	4km		3km	6km	9km
Santa Catarina	22km	27km	26km	27km	8km	3km		3km	6km
Santo Tomas Chautla	22km	28km	26km	28km	8km	6km	3km		4km
Art. Primero Constitucional	24km	29km	30km	31km	12km	9km	6km	4km	

Considering that each of these communities is a demand point (the nodes in a network), we have to find the optimal location for the new fire station that will provide the service. Due to the low population density in the communities of set 2, the installation of a fire station at any community within this Set not justified. Thus, the fire station facility to install should be assigned to a candidate site in Set 1 and must try to maximize the coverage demand of all the other communities within Set 1 and Set 2. Also, San Baltazar Tetela community belonging to Set 1 cannot be considered as candidate because of its geographical location (exposed to floods).

The data shown in Table 3 will be the input for the Maximum Coverage Model, being *i* the set of demand nodes (9 communities: a, b, c, d, e, f, g, h, i), *j* the set of candidate facility sites (Only 4 out of the 5 communities from Set 1: a, b, c, d), h_i the demand at node i, P= number of facilities to locate, and the following parameters:

$$Z_i = \begin{cases} 1 \text{ if node } i \text{ is covered} \\ 0 \text{ if not} \end{cases}$$

$$X_j = \begin{cases} 1 \text{ If we locate facility at candidate site } j \\ 0 \text{ if not} \end{cases}$$

$$\alpha_{ij} = \begin{cases} 1 \text{ If demand node i can be covered by a facility at candidate site j} \\ 0 \text{ if not} \end{cases}$$

The model is formulated as follows:

Maximize

$$\sum_i h_i Z_i , \quad (1)$$

subject to:

$$\sum_j \alpha_{ij} X_j \geq Z_i , \quad (2)$$

$$\sum_j X_j = P, \quad (3)$$

where

$$X_j = 0, 1, \quad (4)$$
$$Z_i = 0, 1. \quad (5)$$

The objective function (1) maximizes the number of covered nodes. As mentioned previously for this case we do not take into account demand values because we are trying to maximize the amount of communities that can be covered within the 10 minutes restriction time.

The constraint (2) indicates that demand at node *i* (each of the nine communities) cannot be considered as satisfied unless at least one of the facilities assigned j (any of the 4 communities of set 1), is selected to cover it. This condition of satisfied demand is based on the 10km restriction (taken from data of Table 3), if the demand point *i* is located within the 10km of the assigned j site, then it can be considered as satisfied. In this same restriction, α_{ij} is a binary variable defined by the capacity of facility assigned *j* to satisfy demand at *i*.

Second restriction (3) ensures that the number of facilities to place have a limit (P), for this case P=1 due to budget limitations, while restriction 4 and 5 indicate that variables X and Z are binaries.

4 Results

Given the input data, LINGO software was used to solve the model. The results showed that under the 10 kilometers restriction, any of the 4 candidates could pro-vide the service for the communities in Set 1, yet, none of them would be able to cover the communities in set 2. As there is still a need to install a fire station that can at least reduce the response times compared to the current attention service, a sensibility analysis was generated to evaluate different scenarios in order to recognize the scope of the service under different distance limits.

The starting distance limit for the tests was the 10 km of the original restrictrion. Following 3 tests increased by 5 kilometers. Results of these tests are shown in Table 4 and the analysis indicates that a significative change happens when the coverage extends to 25 kilometers, which is the distance limit where 7 out of the 9 communities can be considered as covered. However it is not until the distance limit is 27 km when all of the communities can be considered as satisfied if the fire station is located at Santa Catarina community.

Despite this is still a high response time (assuming 27km=27 minuts), it is the one that can satisfy all the current communities demand in the fewest time.

Table 4. Results of Lingo for facility location based on distance limit.

Distance Limit	Facility Location Candidate	Observations
10 km	Art. Primero Constitucional Santo Tomas Chautla Santa Catarina San Pedro Zacachimalpa	Following communities can't be covered: San Andres Azumiatla Santa Maria Guadalupe Tecola Los Angeles Tetela El Aguacate
15	Art. Primero Constitucional Santo Tomas Chautla Santa Catarina San Pedro Zacachimalpa	Following communities can't be covered: San Andres Azumiatla Santa Maria Guadalupe Tecola Los Angeles Tetela El Aguacate
20 km	Art. Primero Constitucional Santo Tomas Chautla Santa Catarina San Pedro Zacachimalpa	Following communities can't be covered: San Andres Azumiatla Santa Maria Guadalupe Tecola Los Angeles Tetela El Aguacate
25 km	San Pedro Zacachimalpa	Following communities can't be covered: El Aguacate Santa Maria Guadalupe Tecola
27 km	Santa Catarina	All communities are covered

5 Conclusions

When a public facility location problem arises, the main issue is to provide the proper coverage under a set of restrictions given by many conditions. These conditions and restrictions define the mathematical model to be used. However, an integration of multiple tools and techniques as well as an analysis is required to ensure that all variables that might affect the result are considered.

This research developed an analysis of current fire stations in Puebla City in order to recognize the coverage service and the need for new fire stations. It was found that due to population growth, some communities settled down in areas within the city that might not have a proper coverage, therefore a new fire station would have to be installed to provide the service. The characteristics of the problem indicated that the Maximal Covering Model was the most suitable to find the optimal location for the new facility. The results showed different scopes of coverage based on the distance limit. Although for emergency services, the fewer the time the fewer the risk, it is important to consider that we are looking for equilibrium, meaning to provide the service for as much communities as possible.

It is important to mention that the Fire Department of the Puebla City does not have (or was not allowed to provide due to security issues) a historical record of demand by type and area. That is why, for this research, this variable was not taken into account, however, we consider that having this type of data could have made a considerable change in the results.

References

1. Rich, S., Davis, K.: Geographic Information Systems (GIS) for Facility Management. IFMA Foundation. Manhattan Software (2010)
2. Mosquera, P., Jennis, C.: Optimal areas determination for new fire stations located at Bogota based on special analysis tools. Universidad Militar Nueva Granada (2015)
3. Wei Gu, Xin Wan, S., McGregor, E.: Optimization of preventive health care facility locations. International Journal of Health Geographics (2010)
4. Church, R.L., ReVelle, C.: Maximal covering location problem. Papers of the Regional Science Association, 32 (1), pp. 101–118 (1974)
5. Coordinación Nacional de Protección Civil, CENAPRED: Atlas Nacional de Riesgos. Indicadores Municipales de Peligro, Exposición y Vulnerabilidad. Accessed June 17, 2016, http://www.atlasnacionalderiesgos.gob.mx/archivo/indicadores-municipales.html
6. Commission on Fire Accreditation, International Template: Standards of Response Coverage. Accessed June 25, https://www.iafc.org/files/1VCOS/sop_CPSE_CFAI_Standard_of_Cover_Template.pdf
7. Daskin, M.S.: Network and discrete location: Models, Algorithms and Applications. John Wiley & Sons, Inc., New York, 520 (1995)
8. Vitoriano, B.: Modelos Matemáticos de ayuda a la decisión en Logística Humanitaria. Depto. Estadística e Investigación Operativa I, Universidad Complutense de Madrid (2012)
9. Buzai, G.: Modelos de localización-asignación aplicados a servicios públicos urbanos: Análisis espacial de Centros de Atención Primaria de Salud en la ciudad de Bogotá (2011)
10. Owen, S.H., Daskin, M.S.: Strategic facility location: A review. European Journal of Operational Research, pp. 423–447 (1998)
11. Parvene, Z., Seyedeh, Z.: Analysis of the Spatial distribution and Locating fire stations, city Bracnch using GIS. Accessed July 29, 2016, ww.euromed.uk.com/files/documents/M-105.pdf (2015)

ISSN 1870-4069

Operational Risk Assessment in 3PL for Maritime Transportation

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

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

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

{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. Operational risk is one of the most important risks in supply chains. The operational risk assessment process includes risk identification and evaluation and prioritization. On the other hand, 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. Maritime transportation is a fundamental activity in global supply chain and it is essential for the commercial trade. We propose a multicriteria approach for risk assessment in a 3PL company for maritime transportation. The multicriteria model uses Fuzzy QFD for risk evaluation and prioritization. The approach is applied in an international company with a branch office located in Colombia.

Keywords: supply chain risk management, operational risk, multicriteria, fuzzy QFD, maritime transportation.

1 Introduction

Maritime transportation is one of the most important industries with its immense share in the global trade. Maritime transportation is a cost-effective method which enables companies to transfer an international cargo between two seaports [1].

The increasing need of the companies to focus on the core business object has generated a trend oriented to outsource different activities. In this context, supply chain activities have evolved from a first stage where we have companies that are responsible for their logistics processes up to the current trend with companies who have delegated all their logistics activities to specialized agents [2].

According with this, maritime companies are important 3PL organizations and have a direct influence in the strategic results of the supply chains therefore managing risk in these companies is an important activity for the supply chain risk management.

In recent years, fuzzy QFD has become a widely used quality tool developed to satisfy customer need in product design and development. Fuzzy QFD provides a mean of translating customer requirements into appropriate technical requirement for each stage of product development and production [3] but in recent years its scope has been expanded towards multicriteria decision making.

The basis of QFD is to obtain and translate customer needs into engineering characteristics, and subsequently into part characteristics, process plans and production requirements [1].

Although there are many papers related with QFD and FQFD applications in maritime transportation ([4][5][6][1][7][8][9]), none of them applied FQFD to operational risk assessment.

Besides, there are some papers related to QFD applications in risk management such as [10] and [11] still there is a gap in risk assessment research and Fuzzy QFD.

QFD approach to enhance maritime supply chain resilience taking both customer requirements and maritime risks into consideration is presented in [4] however still is necessary to include risk prioritization which is the aim of this research. Table 1 shows the scope of some papers related to QFD and FQFD applications in supply chain, 3PL and 4PL, risk and maritime transportation.

Table 1. Papers related to QFD and FQFD applications.

Applications	Papers
QFD and risk management	[11][10]
FQFD in supply chain management	[12][13][14][15][16][17]
FQFD in 3PL (4PL) applications	[18][19]
QFD in maritime transportation	[5] [6][1]
FQFD in maritime transportation	[7][8][9]
QFD in maritime transportation and risk	[4]

We proposed a novel approach to risk assessment using FQFD to prioritize the risk according to the strategical objectives of the company. The approach is presented in the next section and we show an application in a real case in Colombia.

2 Methodological Approach

We proposed a FQFD methodology. This proposal is based in [12] and Fig. 1 shows its development. Previous to phase 1, it is necessary to define the team (experts of the company that are directly involved in the decision process).

The first phase is to stablish the internal variables "WHATs" and then defining the relative importance according to the linguistic scale presented in Table 2. This scale can change if the company or the experts like to use another; the important is that the scale involves linguistic variables and their correspondent fuzzy numbers.

Once the internal variables are defined the team must identify the strategic objectives (HOWs) related to the process where the risks are considering.

To determine the weight of the "HOWs" in the phase 4 the team must stablish the "WHAT" – "HOW" correlation scores using the scale in Table 2.

Finally, the experts determining the risks impact on the strategical objectives and obtain the risk priority. According to this priority, the company can stablish the mitigation or elimination plans to improve the operational risk management process.

Table 2. Linguistic scale to FQFD.

Linguistic variable	Fuzzy number
Very Low (VL)	(0, 1, 2)
Low (L)	(2, 3, 4)
Medium (M)	(4, 5, 6)
High (H)	(6, 7, 8)
Very High (VH)	(8, 9, 10)

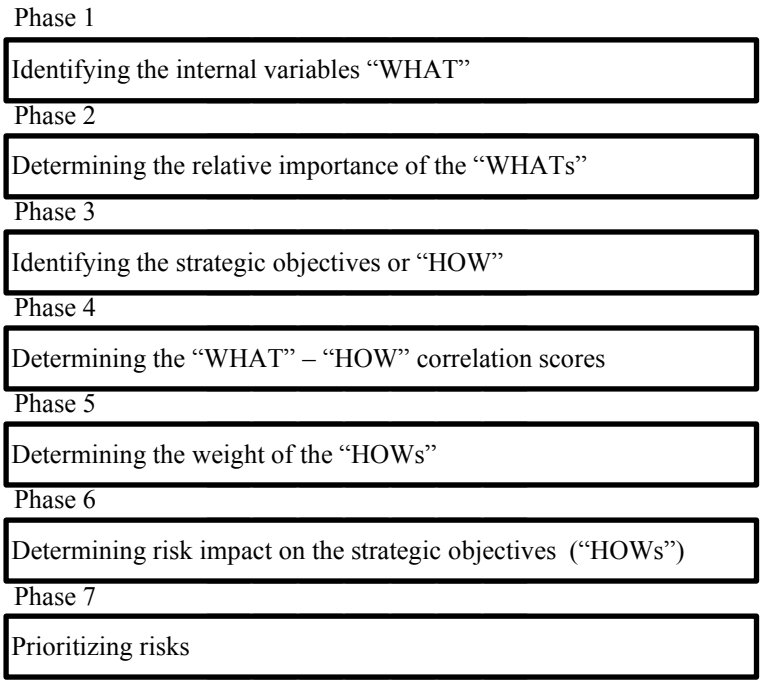


Fig. 1. Methodological approach to risk prioritization.

3 Case of Study

The methodological approach was used in a maritime company with filial in Colombia. This company is one of the largest and the most recognized in global container shipping.

The company operates 480 offices in 150 countries. It owns 480 ships and operates 200 routs in the world.

The team was composed by the operations manager (OM), commercial manager (CM), export manager (EM) and branch manager (BM). According to the methodological approach the first phase was to define the internal variables "WHATs" and their weights in triangular fuzzy numbers (TNF) (Table 3).

Table 2. WHATs and their weights.

WHATs	OM	CM	EM	BM	Ponderations (TFN)		
Good resource utilization - **GRU**	H	VH	VH	H	7	8	9
Good customer service – **GCS**	VH	H	H	H	7	8	9
Economic benefits – **EB**	VH	VH	VH	VH	8	9	10
Company positioning –**CP**	H	M	H	M	5	6	7
Market participation increasing – **MPI**	M	H	H	VH	6	7	8

The next phase was to identifying strategical objectives (HOWs), in this case they were: Improve market share position, Optimize profit per shipment, Ensure market coverage, Optimize space ships and Customer loyalty

Then the team determines the "WHAT" – "HOW" correlation scores. In Fig. 2 these correlations are presented. For example the team opinions about the impact of improve market share position with the good resource utilization is high for the operations manager, commercial manager and export manager, but is medium for the branch manager

	Strategic objectives (HOWs)																			
WHATs	Improve market share position				Optimize profit per shipment				Ensure market coverage				Optimize space ships				Customer loyalty			
	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM
GRU	H	H	H	M	VH	H	H	H	VH	M	H	VH	H	VH	H	VH	H	M	M	H
GCS	H	H	VH	H	H	M	M	H	VH	M	M	H	H	M	M	M	VH	VH	H	H
EB	VH	H	VH	VH	VH	H	H	VH	M	H	VH	M	VH	H	H	H	H	VH	H	H
CP	H	H	VH	H	M	M	L	L	VH	H	H	H	H	M	M	M	VH	H	M	H
MPI	VH	VH	H	H	M	M	H	H	VH	H	H	H	L	M	M	L	VH	H	M	H

Fig. 2. "WHAT" – "HOW" correlation scores.

The next phase consists in determining the risks impact in the strategical objectives.The risks were identified previously and they are presented in table 4. The Fig 3 shows the risks impacts and finally in the Table 5 prioritization of risks is presented.

The results show that the most important risk is low availability of spaces and the company must develop plans to mitigate or eliminate it to improve the supply chain performance. The second risk in priority is low availability of containers and although

lack of documents or instructions is the low risk once the others have implemented actions also must be addressed.

Table 4. Risks identified to maritime transportation (container shipping).

Abbreviation	Risks
LAS	Low availability of ships
DDD	Delay in delivery of the documents
FG	Freight contamination
LDI	Lack of documents or instructions
MI	Mistakes in information
CSS	Changes in services and schedules
LAS	Low availability of spaces
LAC	Low availability of containers

	Improve market share position				Optimize profit per shipment				Ensure market coverage				Optimize space ships				Customer loyalty			
	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM	OM	CM	EM	BM
LAS	H	H	VH	M	VL	L	L	VL	L	L	M	VL	VH	VH	H	VH	L	M	M	L
DDD	H	M	M	M	L	L	VL	VL	M	M	H	M	VH	H	VH	VH	VH	H	M	H
FG	L	L	VL	L	VL	L	L	L	L	L	L	L	H	H	VH	M	VH	M	H	H
LDI	M	H	L	M	L	L	L	VL	M	L	L	VL	L	M	VL	L	L	M	VL	L
MI	M	H	L	M	L	L	L	L	M	L	L	VL	H	H	VH	H	H	M	M	L
CSS	VH	VH	VH	H	VH	H	H	M	VH	H	M	H	H	H	H	H	H	VH	M	H
LAS	VH	VH	VH	VH	VH	H	VH	VH	VH	H	M	H	VH	VH	VH	H	H	H	VH	M
LAC	VH	H	VH	VH	H	VH	VH	VH	H	H	H	H	VH	M	H	H	H	H	M	VH

Fig. 3. Risks impact in strategic objectives.

Table 5. Risks prioritization.

Risk	Triangular fuzzy number			Crisp number
Low availability of spaces	268	412	602	**424**
Low availability of containers	253	393	576	**404**
Changes in services and schedules	242	378	557	**389**
Delay in delivery of the documents	180	294	448	**304**
Low availability of ships	149	252	394	**262**
Mistakes in information	140	241	379	**250**
Freight contamination	128	224	358	**234**
Lack of documents or instructions	90	173	290	**181**

4 Conclusions

We present a novel approach to prioritize risks in supply chain activities involving third party logistics companies (3PL) and particularly in maritime transportation. This approach allows defining the most important risks according with the strategical objectives of the company and thus the company can establish action plans to mitigate or eliminate the risks.

Although in literature there are some papers that using QFD in risk management, there is a lack of papers using FQFD for risk prioritization. Our proposal aims to encourage more works in this field.

We showed that FQFD is a valuable tool in the operational risks assessment so it is important to continue exploring the usefulness of this tool in operational risk management system.

We can prioritize risks according to the strategical objectives of the company and this impact directly in its results and with these results it is possible for the company to define actions in line with its strategy to risk management.

The methodology should be applied throughout the supply chain in order to improve its overall performance.

References

1. Sener, Z., Ozturk, E.: A QFD-Based Decision Model for Ship Selection in Maritime Transportation. Int J Innov Manag Technol, 6, pp. 202–205, doi: 10.7763/IJIMT.2015.V6.602 (2015)
2. Manotas, D.F., Osorio, J.C., Rivera, L.: Operational Risk Management in Third Party Logistics (3PL). Handb Res Manag Strateg Achiev Optim Perform Ind Process, doi: 10.4018/978-1-5225-0130-5(2016)
3. Dat, L.Q., Phuong, T.T., Kao, H.P., et al.: A new integrated fuzzy QFD approach for market segments evaluation and selection. Appl Math Model, 39, pp. 3653–3665, doi: 10.1016/j.apm.2014.11.051 (2015)
4. Lam, J.S.L., Bai, X.: A quality function deployment approach to improve maritime supply chain resilience. Transp Res Part E Logist Transp Rev., doi: 10.1016/j.tre.2016.01.012 (2015)
5. Lam, J.S.L.: Designing a sustainable maritime supply chain: A hybrid QFD-ANP approach. Transp Res Part E Logist Transp Rev., 78, pp. 70–81, doi: 10.1016/j.tre.2014.10.003 (2015)
6. Razik, M.A., Tahar, R.M., Hasrulnizzam, W., et al.: Integrated Quality Function Deployment (QFD) Model for Dry Bulk Terminal Improvements (DBTI) in Malaysian Ports. J Econ Bus Manag, 3, pp. 4–7, doi: 10.7763/JOEBM.2015.V3.219 (2015)
7. Liang, G.S., Ding, J.F., Wang, C.K.: Applying fuzzy quality function deployment to prioritize solutions of knowledge management for an international port in Taiwan. Knowledge-Based Syst, 33, pp. 83–91, doi: 10.1016/j.knosys.2012.03.012 (2012)
8. Ding, J.F.: Applying fuzzy quality function deployment (QFD) to identify solutions of service delivery system for port of Kaohsiung. Qual Quant, 43, pp. 553–570, doi: 10.1007/s11135-007-9138-7 (2009)
9. Celik, M., Cebi, S., Kahraman, C., Er ID.: An integrated fuzzy QFD model proposal on routing of shipping investment decisions in crude oil tanker market. Expert Syst Appl., 36, pp. 6227–6235, doi: 10.1016/j.eswa.2008.07.031 (2009)

10. Gento, A.M., Minambres, M.D., Redondo, A., Perez, M.E.: QFD application in a service environment: A new approach in risk management in an university. Oper Res, 1, pp. 115–132, doi: 10.1007/BF02936289 (2001)
11. Costantino, F., De Minicis, M., González-Prida, V., Crespo, A.: On the use of quality function deployment (QFD) for the identification of risks associated to warranty programs, 6, pp. 4440–4449 (2012)
12. Bevilacqua, M., Ciarapica, F.E., Giacchetta, G.: A fuzzy-QFD approach to supplier selection. J Purch Supply Manag, 12, pp. 14–27, doi: 10.1016/j.pursup.2006.02.001 (2006)
13. Bottani, E., Rizzi, A.: Strategic management of logistics service: A fuzzy QFD approach. Int J Prod Econ.,103, pp. 585–599, doi: 10.1016/j.ijpe.2005.11.006 (2006)
14. Sohn, S.Y., Choi, I.S.: Fuzzy QFD for supply chain management with reliability consideration. Reliab Eng Syst Saf., 72, pp. 327–334, doi: 10.1016/S0951-8320(01)00022-9 (2001)
15. Zarei, M., Fakhrzad, M.B., Jamali, M.: Food supply chain leanness using a developed QFD model. J Food Eng., 102, pp. 25–33, doi: 10.1016/j.jfoodeng.2010.07.026 (2011)
16. Hassanzadeh, S., Razmi, J.: An integrated fuzzy model for supplier management : A case study of ISP selection and evaluation. Expert Syst Appl., 36, pp. 8639–8648, doi: 10.1016/j.eswa.2008.10.012 (2009)
17. Kazançoğlu, Y., Aksoy, M.: A fuzzy logic-based quality function deployment for selection of e-learning provider. Turkish Online J Educ Technol., 10, pp. 39–45 (2011)
18. Leina, Z., Tiejun, P., Guoqing, Y.: Process Control and Simulation for the Fourth Party Logistics Using Value Chain Analysis. 2010 Int Conf Electr Control Eng., pp. 1503–1506, doi: 10.1109/iCECE.2010.371 (2010)
19. Wang, F., Li, X.H., Rui, W.N., Zhang, Y.: A fuzzy QFD-based method for customizing positioning of logistics Service products of 3PLS. 2007 Int Conf Wirel Commun Netw Mob Comput WiCOM, pp. 3326–3329, doi: 10.1109/WICOM.2007.825 (2007)

ISSN 1870-4069

Inventory Reduction in a Metal-Mechanical Industry Using Logistical Tools

Alejandro H. Martínez-Márquez, Diana Sánchez-Partida, José Luis Martínez-Flores

Universidad Popular Autónoma del Estado de Puebla,
Puebla, México

alejandroheriberto.martinez@upaep.edu.mx, diana.sanchez@upaep.mx,
joseluis.martinez01@upaep.mx

Abstract. This paper seeks the critical factors for a correct choice in the components set up along the supply chain. ABC classification was made for all the suppliers according to the supplied items for the company. After a demand analysis, the seasonal nature and trend during the timeline, we used the last four years registers where is shown a decreasing behavior in 30% of the items; Economic Order Quantity (EOQ), for deterministic (<20%) which 58 items represents 64% of the A classified items into the pull system; and a continuous review model q, R for stochastic (>20%) which 33 items represents 67%, 429 items (92%) are classified as B and 2,128 items (99%) are C. We reduced the lot quantity orders in 32%, the re-order points were adjusted according to the EOQ and q,R models; Safety stock levels were reduced 55% representing $450,000 pesos against the initial inventory of 1'000,000 pesos.

1 Introduction

Nowadays more companies acknowledge the importance of stock as a key component for business, even though its role was previously underestimated. Nowadays, stocks are recognized as a main component within the supply chain and other commercial procedures. According to Aznar (1986), the role of stocks in business is increasingly gaining in importance:

For an industrial company, between 25% and 30% of the total assets is represented by the row of inventories. Yet, for commercial the financial statements can reach up to 80% or even a higher percentage. Commonly, part of the resources invested in inventories comes from loans to other companies (Aznar, 1986) representing a high cost impact.

Inventories concentrate the totality of assets that companies require to satisfy the demand of products or services that are offered. The main reasons to justify the maintenance of inventories is to reach economies of scale, to face uncertainty and speculation, to leave out temporal Investment, to smooth the variability of demand, the restrictions of business logistics and the cost of control systems (Nahmias, 2007).

According to Heizer and Render (2001) the costs associated to the inventories can be classified in two categories: ordering costs and maintenance costs. The ordering costs include the specifications, purchase order, suppliers follow-up, office documents

and administrative staff which are needed to process the orders. The maintenance cost include; the warehouse rental (if it is the case) depreciation, operational costs, taxes, cost of capital opportunity, insurances, financing costs, financial losses, waste, obsolescence, etc. Through this method, the acquisition of materials has to be done using the Just in Time (JIT). However it is not pure JIT since it combines tools from other systems, it is a hybrid version of JIT that fill the gaps that cannot be completed due to an array of deficiencies in the supply chain.

This project is aimed at the personalized products and high volume business unit. It focuses on analyzing -from a single business unit- the 24 thousand part numbers from the stock of direct components for assembly. The objective is to improve the efficiency of purchase processes, to generate savings for the company and to optimize the inventory levels. In a first instance, the basic components of the business unit in question were downloaded. A Pareto of suppliers was applied in order to identify the main part numbers of the company, where the amount of items designated is displayed. Using this information, the supplier that provides more part numbers was selected and the part numbers were classified according to their profit and use. This classification of the articles was used to analyze the behavior of the demand.

In a later stage, the Coefficient of Variability was calculated by using the information gathered from the part numbers selected in order to foretell the behavior of the demand. The results obtained from this process were used to establish the model to be applied the two cases found. Finally, the safety stock was calculated after the results of lot sizes were concluded.

2 Literary Review

The Pareto principle -also known as the 80-20 rule or the principle of factor sparsity-, named after the Italian economist, sociologist and engineer Vilfredo Pareto who introduced the concept for the first time, is considered one of the major contributions to microeconomics and a baseline for ABC analysis. This concept divides the inventory available in three categories based on the income annual volume.

The Pareto principle indicates to separate 'the vital few from the trivial many'. The aim is to apply policies for inventories so that there is a focus on the vital few parts of an inventory instead of the trivial many. It is not realistic to monitor cheap articles with the same emphasis than the costly ones (Heizer J. & Render B., 2009).

Classifying part numbers in an inventory is of great importance to maintain efficient levels. Dhoka et al (2013) show that in the Analysis ABC -based in the Pareto principle- 20% of articles contribute to 80% of sales. Generally, this 20% is composed by articles classified as A and produce the 80% of the income of the company. In their study they analyze 330 articles from the automotive industry within a specific time frame in order to show the changing behavior of A, B and C categories.

Another tool used is the Economic Order Quantity model or EOQ which allows to determine the amount of items that should be ordered to minimize costs. This model is one of the preferred to achieve a range of demand. It allows to balance the fixed costs per lot and the logistical costs (Roach Bill, 2005). The EOQ is not a complex model, it is based on finding the point where the costs of ordering a product are equal to the costs of maintaining it.

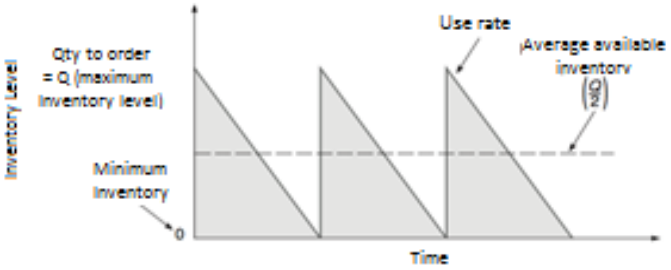


Fig 1. Graphic representation of EOQ model (Heizer, 2009).

The graphic for the inventory use through the time has a dented saw shape, as is illustrated in Fig 1. In the figure Q represents the quantity to be ordered. Generally, when an order is received the inventory level increases from "0" to Q units. Due to the constant demand through the time, the inventory decreases in a constant rate during a specific period of time. Every time the inventory level reaches to zero, a new order is placed and received; and once again the inventory level reaches to Q units (represented by vertical lines. The process continue through the time indefinitely (Heizer, J. 2009).

Sipper, D. & Bulfin, R.L. (1998) determine the q, R model as special for the continuous review system. This model has a unique optimization of service and focus, equivalent to the stochastic version for a deterministic EOQ. Both variables establish the models policy, where Q represents the order size and R is the reorder point. The principal issue in the analyzed company, according to the study, is the low rotation and high inventory level. Qing Li et al (2008) describes the problem of no defined demand/production, through the model definition to calculate the lot size, having all the used variables.

Pratap, M., & Singh, H. (2008), proposed an stochastic inventory model in order to select the economic order to improve the inventory level control, and delete the stock out raw material situation. The result of this study revealed that the companies can save a considerable yearly amount of money adopting the proposed model.

As Mattsson (2007) describes in his study subject, with the showed simulations in his work is a simply way to describe the positive effect of the continuous inventory reviews expressed using equations trying to clarify the model. Below the Figure 2 are showed 2 examples of how the re-order points can be reached, where it is considered a minimal inventory level meanwhile the placed order is triggered to be received during a specific period of time.

The main issue that can be represented in this model, is the stock out for raw material, and is considered in its critical point when the replenishment has not been received, in other words, meanwhile the material is under production with the supplier or in-transit to the needed facilities. For these reasons the lead time plays an important role inside the company's supply chain, creating several ideas to encourage a deep and strong partnership between supplier-client as a single commercial society.

Hopp, Spearman & Zhang, in its publication Easily Implementable Inventory Control Policies (1997), when customer service depends mainly in the material availability, the problem has a solving estimation with the minimization for the total inventory investment subject to average fill rate restrictions and orders frequency. In

which the authors selected the q, R model as the most appropriated playing the individual inventory role.

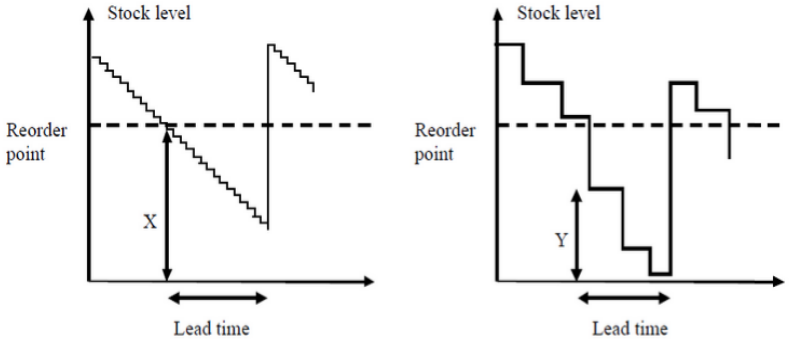


Fig 2. Re-order point behavior example.

Hadley & Whitin (1963) formalized the q, R model in a costs framework, where the total amount of costs is composed by ordering, maintaining and backordering costs. Federgruen & Zheng presented a simple and efficient algorithm to calculate a q, R optimal lineal policy in Q*; Zheng (1992) contributed with important ideas for the heuristic showing the utilization of EOQ formula instead of the optimal order quantity, Q*, finding the optimal order point, r, causes a cost increase for the order in a rate of no more than 1/8. Also demonstrated when the saving in inventory is reflected, the relative total cost increase is much lower than in the worst scenario.

Priyan y Uthayakumar (2015) presented that even that the continuous review inventory model has been studied for several years, helping to minimize the related costs to the ordering quantities, lost sales and process order costs, exists some paradigms into the inventory areas. The most common is the assumption that received quantities are equal to the ordered quantities; which is a non-complete true asseveration due to the huge amount of variables during the process, packing and delivery moments. For all this reasons they developed a mathematical model including the stock out costs, these are located in the studied company we are studying.

3 Problem Description

The selected company is dedicated to manufacture mechanical seals in Tlaxcala. It has been running for more than 30 years and nowadays has 400 employees working in a 3 shifts scheme. Reviewing the general acquisition of components process we can highlight the suppliers' ship daily via FedEx / DHL due to the characteristics of size and weight of the part numbers. 70% of the suppliers are located in U.S. 20% are in Europe and 10 % between Asia and Middle East. Actually the use the financial strategy of non-pure Just in Time (JIT), because of the always changing characteristics of the business, they have incorporated other strategies like the Vendor Management Inventory (VMI).

The productive strategy is Make-to-Order, as they start producing until a sales order is entered and reflected into the ERP system. There are some parallel running processes with sales where drawings are developed, material specifications and creation of new codes when applies. This will help the final customer receive their material in the compromised lead time.

The company use an old fashioned ERP to follow up the open orders having an operational control, creating a data base visible for the other sites around the world. This tool was developed by IBM and have his boom during 90's. Nowadays it is still working in a proper way for the business requirements. The last massive parameterization was performed 5 years ago. However the buyers review individually their own part numbers before placing a new order. The business have a consecutive growth for the last 5 years, but we identified 770 part numbers (30% approx.) with a decreasing usage because the engineering reviews do not need those pieces anymore. New business have been incorporated and calls for components with specific restrictions as FDA's compliance. As consequence new part numbers have been added to the warehouse stock ignoring the ABC classification; the last time they reviewed their safety stock levels was done in 2012. These are the mainly reasons why we cannot define if they have a overstock level, there are many logistical costs over the sky and an important lack of raw material acquisition planning, giving as a result penalties for stock outs and re-works for the rest of the involved areas with negative impact on the financial metrics.

The actual needs of the market demands less lead time or deliveries and improvement for the competitive prices with no quality issues. Other companies that produce the same type of seals are increasing the competences in the critical involved areas (Purchasing, Warehouse and Finances) where the major part of the expenses are reflected.

As a first stage an ABC classification was performed for the 144 suppliers with active orders during 2015, according to the part numbers volume supplied to the company. The result is that only one vendor has assigned the 33% of all the existing part numbers in loaded in the system. With the selected supplier, we worked in the second stage where an ABC classification was done for 2,636 items; using a coefficient of variation and pattern demand to establish the most accurate model to be used. If the coefficient of variation is less than 20% a deterministic model EOQ will be used, that determine the lot sizes for the economic order. Otherwise, if the value is higher than 20%, the best option will be working with a q, R stochastic model, that will define the size of the order and re-order point with irregular pattern demand.

Due to the uncertainty in the demand, as the company works under personalized projects, a safety stock is needed only over the critical part numbers that can be used in multiple seals as a common base with long lead time enough to impact the customer order. During the third and last stage, the Safety stock was analyzed against the information loaded in the system with the intention of update the purchasing parameters. The implementation team was conformed of 3 departments Warehouse, Purchasing and Finances.

4 Methodology and Results

4.1 ABC Classification for Supplier and Components

There are 24,059 items loaded into the system; only 7,879 items had movement during the last year. In order to narrow the scope a Pareto diagram was made where the assigned part numbers per vendor are shown, in Figure 3 A, B, C,… etc. represents the supplier name. In total we have 144 active suppliers delivering items, but only 20 of them have assigned the 80% of the total amount of part numbers.

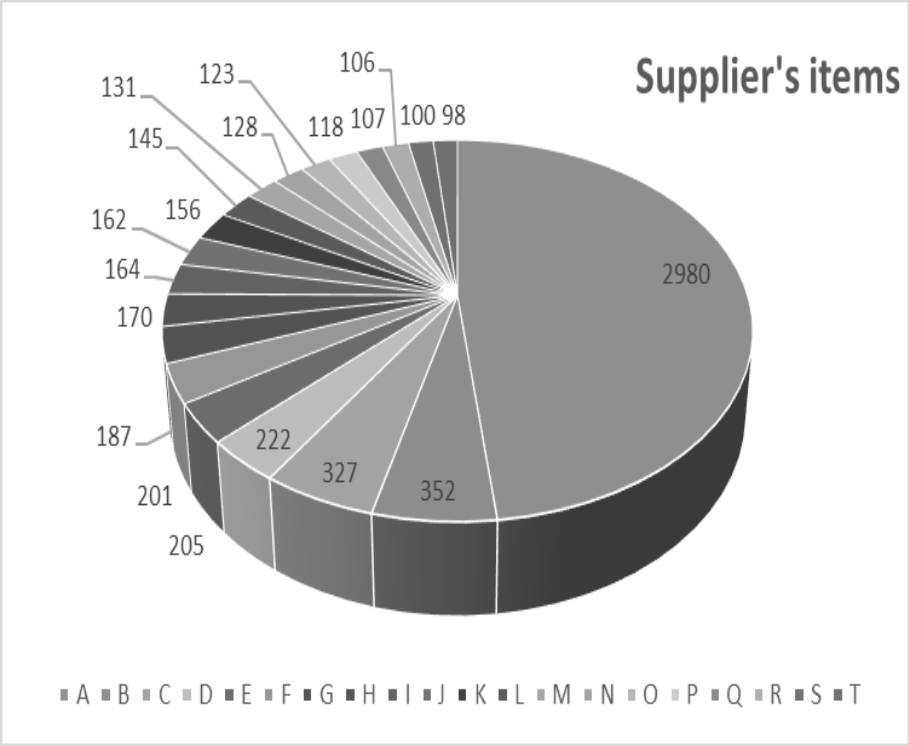


Fig 3. Items assigned per supplier.

The "A" supplier has 2,980 assigned items, because of that we will continue only with this supplier. A modification in that quantity of items will not only have effects of inventory reduction, a drastic logistical costs reduction as the inventory models explains.

Taking 2,980 items, we proceed with ABC classification as shown in the following charts; Profit based on (Chart 1a) and Usage based on (Chart 1b) with 2015 database to have an assertive result.

In the second stage we considered that the first simple of 91 items is not enough to specify a method, not even the model to use with the other part numbers. We took the decision to open the sample to 100%, so all the part numbers assigned to the supplier will be reviewed.

PROFIT BASED CLASSIFICATION		
A	B	C
701561-Q1	K23M1875GX3	568138AF
393985PS3	568141WG56	4N02100CK
MLB32752919	MLA37358225	MEAX8840252
C0164845YO	KU3R1125333	4R0429NL
TCN3029	4N01736NL	621533DB
3R26481YO	4R13467DB	568112AF
52349 90 1 1-Q1	KCUM2250R33	Y53A5500AC3
KU3R1750333	568333ZV	Y5FQ3375M33
TCN2624	KR8G2750G33	MYCG5875AJ3
3R98758K30	KR3D2625FA3	670719DB
TCN2764	TIRM2000X33	614608GE
B0235952-Q1	MLB33460633	MKAX88402BE
B0246289-Q1	668815LV	MCA23280714
MLB32752906	MEAX8840239	4N02099CK
49226 09 1 2-Q1	4N01730AI	MIAX8840250
48365 04 1 2-Q1	3N49141DB	618947MG
AU3J1875MZ3	KIRM2250TK3	626671LP
MLB32752920	TCRM1500X33	568357AF
KU3C1875C33	568234GU	4N11498GE
4R0440X902	4R2893NL	K43R1000333
Total	*Total*	*Total*
526	*747*	*1683*

USAGE BASED CLASSIFICATION		
A	B	C
MLA37358228	4R39511DB	MNA20442233
C0012107NL	4R3483CI	UR437CK
C0012112X974	4N02083X409	568150MG
C0195336ZZ	MYCG5115AB3	MYCD1375CA3
MLA21480433	4R15512DB	568222MG
MKAX8840217	C0013644NL	4R2888CI
MKAX8840212	4R2951CI	K23D3000Q33
MLA37358225	4R3269CI	156407GE
668855DB	568031MG	664282DB
C0195337ZZ	4R45404DB	4R0437HU
668802DB	4R3148ZZ	US325CK
MKAX8840254	MIAX8840250	4N01767DB
C0012111X974	MKAX8840227	0700142330 0
MKAX8840209	4R0457DY	4N01731CI
MKAX8840215	668853DB	4SIPG711X
C0201666ZZ	670453DB	616267GU
4R0434CI	C0013641CK	UR212CK
670425DB	MKAX88402CA	568339MG
C0012118X974	C0013647X974	568226MG
UR243CK	4R3150ZZ	4SIPG714
Total	*Total*	*Total*
91	*464*	*2425*

Chart 1a. Profit.

Chart 1b. Usage.

4.2 Coefficient of Variability

This operation is needed to determine if the used model will be deterministic or probabilistic; in order to take the best decision, the following steps are recommended:

1) Calculate estimated average demand d for each period:

$$\bar{d} = \frac{1}{n}\sum_{1=l}^{i=n} di. \tag{1}$$

2) Calculate estimated variance D for each period:

$$Est.\,var\ D = \frac{1}{n}\sum_{1=l}^{i=n} di^2 - \overline{d^2}. \tag{2}$$

3) Calculate estimated relative demand variation (also called as coefficient of variation). The final amount is represented as CV:

$$VC = \frac{est.\ var\ D}{\bar{d}^2}. \tag{3}$$

If the coefficient is lower than 0.20 (20%) a deterministic model can be used; otherwise, the demand is too instable and a probabilistic model must be used. After calculate the coefficient of variation for all the classified A, B, C items we can observe the behavior in Chart 2a showed below*:*

A			B			C		
%	No. Items	Pattern	%	No. Items	Pattern	%	No. Items	Pattern
0.90	81.00	Regular	0.34	157.00	Regular	0.02	38.00	Regular
0.10	10.00	Irregular	0.66	307.00	Irregular	0.98	2,387.00	Irregular
	91.00			464.00			2,425.00	
%	No. Items	Model	%	No. Items	Model	%	No. Items	Model
0.64	58.00	Deterministic	0.08	35.00	Deterministic	0.003	7.00	Deterministic
0.37	33.00	Probabilistic	0.92	429.00	Probabilistic	0.997	2,418.00	Probabilistic
	91.00			464.00			2,425.00	

Chart 2a. Behavior for ABC classification according to Coefficient of variation.

We can appreciate the part numbers classified as "A", most of them have a regular pattern and a deterministic method can be applied due to their coefficient of variation. Not in the same way for the rest of the items; the scene changes for the B and C classified items; where we can see most of the items have an irregular pattern and it is no possible to apply the same method that is why we call for a stochastic one.

ABC			
Pattern	No. Items	%	
Regular	276.00	9%	
Irregular	2,704.00	91%	
Total	2,980.00		
Model	No. Items	%	
Deterministic	100.00	3%	***EOQ***
Probabilistico	2,880.00	97%	***QR***
Total	2,980.00		

Chart 2b. Global behavior and method to be used according to Coefficient of variation.

Into the *Chart 2b* shows that the EOQ will be used for the 3% of the items with the coefficient of variation acceptable for a deterministic method. For the rest of the items B & C mostly we decided to use the continuous review q, R model.

4.3 Economic Order Quantity (EOQ)

In this stage of the Project we realized the economic lot estimated using the general equation for EOQ model (4) as is showed below:

$$Q* = \sqrt{\frac{2DS}{H}}, \tag{4}$$

where:

D = Yearly unitary demand.

S = Ordering cost for each requirement.

H = Holding monetary cost.

The total logistical cost (5) is obtained summarizing the order cost plus the acquisition cost plus the holding inventory cost:

$$\mu(q) = \frac{kd}{q} + cd + \frac{hq}{2}. \quad (5)$$

The frequency (6) for setting a new order is:

$$T^* = \frac{q^*}{d}. \quad (6)$$

The order number (7) is given by the consume (demand) between the optimal quantity to be ordered:

$$n = (d/q^*). \quad (7)$$

As a result we have a total logistical cost of $4'269,694 pesos; due to the order cost is low, the optimal quantity to order is high; and this is reflected in holding inventory for long periods of time. Also we obtained an average of 4.2 orders per year for almost all the part numbers. Inside the Chart 3 we can see the relation between the part numbers according to their classification and the percentile of the logistical cost that represents for the 100 reviewed pieces.

	# Items	Logistical Cost	%
A	58	$2,386,249.24	56%
B	35	$1,845,183.93	43%
C	7	$38,260.85	1%
Total	**100**	**$4,269,694.02**	

Chart 3. Total logistical cost for EOQ model.

4.4 Continuous Review Model with Uncertain Demand (q, R)

For this project we used the q, R inventory model with a random demand. After setting the needed equations we will have an optimal value for the limits approach, without leaving the order limits allowed for the inventory value, where the theorems will prove the supposed changes in the following orders. It is taken this analysis because we proposed a periodic review with emergency safety replenishment. Characterized with short order periods, absorbing the order cost. Below we can see the equations for the results obtainment.

The standard deviation (8) is calculated during the lead time, multiplying the demand's standard deviation with square root of lead time:

$$S'_d = S_d\sqrt{L} \quad (8)$$

The re-order point (9) is the result of multiplying the mensual demand with the lead time expressed in months plus the adjusted standard deviation with the Z value to the normal distribution with a customer's service level previously defined:

$$R = d * L + z_{CSL} * S'_d. \quad (9)$$

The order size (10) tells us we should calculate the square root of two times the order cost multiplied with the monthly demand between the holding inventory costs:

$$q = \sqrt{\frac{2kd}{h}}. \tag{10}$$

The ratio of the standard deviation (11) stays out or do not complies with the demand 1-s´d*Ez/q gives us the number of non-supplied pieces.

The stock outs ratio given a Z.

Fs (z) = Accumulated normal standard distribution function.

Fs (z) = Density of normal standard distribution function. This is the point where occurs the probability:

$$E_z = z * [F_s(z) - 1] + f_s(z). \tag{11}$$

The total logistical cost (12) $[\mu(q^*)]$ is given by the following equation:

$$\mu(q) = \frac{kd}{q} + h\left[\frac{q}{2} + z_{CSL} * S'_d\right] cd + \frac{d}{q}(u * s'_d * E_2). \tag{12}$$

FR (13) is defined per the following equation:

$$FR = 1 - \frac{\left(\frac{d}{q}\right) * S'_d * E_z}{d} = 1 - \frac{S'_d * E_z}{d}. \tag{13}$$

As a result we obtained a total logistical cost of $5'861,794 pesos. The q, R model considers a cyclical service level of 100% and suggests a re-order point with a lot size for each part number. See the Chart 4 to review the relation between the articles with the logistical costs and the ratio that represents each category.

	# Items	Logistical Cost	%
A	33	$34,152.06	1%
B	429	$539,495.40	9%
C	2418	$5,288,147.06	90%
Total	**2880**	**$5,861,794.52**	

Chart 4. Total logistical cost for q,R model.

4.5 Safety Stock (SS)

The safety stock inventory is an additional quantity for an item inside your warehouse with the objective of reduce the production stock outs. There are additional associated costs that must be considered during the calculation; however the holding inventory cost can be lower than the cost for losing a client if the delivery is not completed timely and properly.

We took 920 items with a safety stock inventory loaded with the selected supplier. The 2,060 pending items do not have a safety stock loaded and will be excluded from our analysis. We decided to perform an analysis between the loaded inventories into the system against a proposal of re-ordering points with the q, R model. After the estimations, we analyzed the implementation team to verify the obtained results.

5 Discussion of Results

At the beginning of this project I have no idea where to find a problematic situation inside the organization. I started downloading different reports in order to see if I can get some clues about a malfunction of any area. Suddenly I get trapped in a huge quantity of information that have no sense for me. I have to accept that into the theory the problems looks like if all could be located easily. After several meetings with different areas, the situation of the inventory came out and the opportunity to work in that issue was set.

The part numbers trend through the analyzed period of time gave us a general overview of the behavior of the sales, and the inventory level effects. The Figure 4a, 4b, 4c & 4d shows some examples of one analyzed item and the movement during the last 5 years. The graphics are based in the yearly usage showing the pieces during the period of time mentioned before.

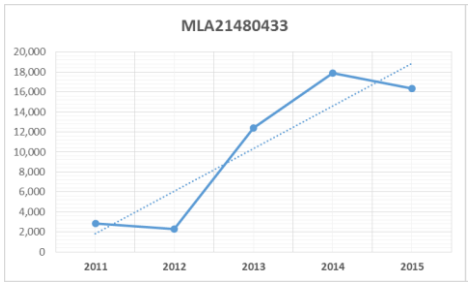


Fig 4a.

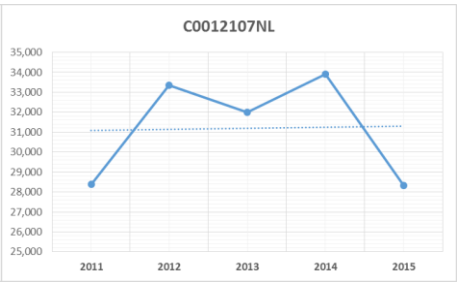


Fig 4b.

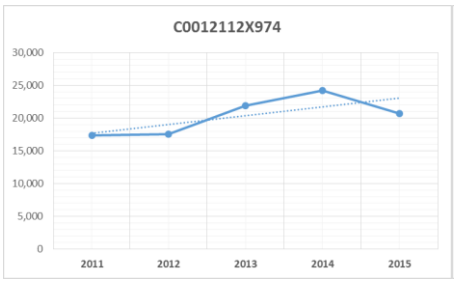


Fig 4c.

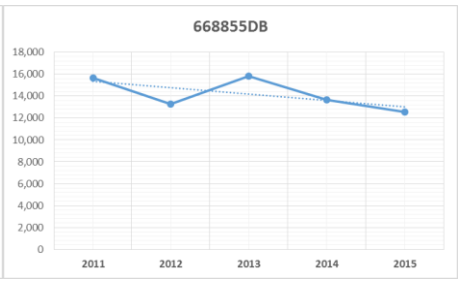


Fig 4d.

Concept	Actual	Model	Difference	%
Items with SS	920	919	-1	0.1%
Inventory value	$1,046,076.00	$996,888.23	-$49,187.77	5%

Chart 5. Actual status against proposed model values.

As we can see into the graphics the behavior of the part numbers are as many as the customers' requests. It is a challenging situation due to the nature of the business, and a great opportunity to solve this issues with the tools available in the science theory.

Nowadays the inventory levels into the companies represents one big opportunity for creating savings without the need additional investments for the operational processes. Every day the highest levels of the organizations increases their awareness about the negative effect of having money stocked into the shelves. In order to accomplish the purpose of this project a selection of logistical tools were used with mathematical operations the values represents the path they are following up.

The inventory value at the beginning of the project was $1'046,076 pesos with 920 part numbers with a safety stock inventory loaded, representing 8% of the total value of the analyzed plant inventory. The modification in the minimal lots suggest 58% reduction of the inventory value and decrease of 691 part numbers in the safety stock items loaded. After changes were applied the saving impact is immediately, even when the available inventory in plant was not considered. This variable was discarded for practical reasons for this project, because we do not have control over the demand and due to the restricted access to the sales department information. Finally the inventory will reach the objective level of $435,854 pesos. Please refer to the Chart 5 to see the initial amount of loaded items with safety stock and their value in pesos, also we can see the difference that represents the change in parameters, where a 5% reduction is showed.

The idea of reduce the inventory level without any investment had a great acceptance, and that is why we reviewed each one of the 920 items, leaving out the obsolete, discontinued or substituted items; also all those special items that only were bought one time without requirement. After the complete review, the implementation team concluded with the Chart 6 values, for reaching an inventory level decreased 58%.

Concept	Actual	Model	Difference	%
Items with SS	920	691	-229	24.9%
Inventory value	$1,046,076.00	$435,845.67	-$610,230.33	58%

Chart 6. Actual state against final agreement.

Chart 7a & 7b shows the inventory value proposed for both cases: q, R model and final agreement with ABC classification. We can see the dramatically decrease in items with C classification and the value increase for A classified items.

q, R Model			
Clasif	**Items**	**Value $**	%
A	83	$63,728.77	6%
B	345	$347,503.09	35%
C	491	585.656.37	59%
Total	**919**	**$411,231.86**	

Chart 7a.

Final Agreement			
Clasif	**Items**	**Value $**	%
A	82	$70,125.72	16%
B	327	$152,672.11	35%
C	282	$213,047.84	49%
Total	**691**	**$435,845.67**	

Chart 7b.

References

1. Aznar, M.: Organización de Almacenes y Control de Inventarios, México: ECASA.P1 (1986)
2. Dhoka, D., Choudary, L.: ABC Classification for Inventory Optimiztion, IOSR Journal of Business and Management (IOSR-JBM), e-ISSN: 2278-487X, p-ISSN: 2319-7668, vol. 15, (Nov. - Dec. 2013), pp. 38–41 (2013)
3. Heizer, J., Render. B.: Administración de Operaciones, México: PEARSON Educación, pp. 978–987 (2009)
4. Hopp, W. J., Spearman, M. L., Zhang, R. Q.: Easily Implementable Inventory Control Policies. Operations Research, pp. 327–340 (1997)
5. Mattson, S.A.: Inventory control in environments with short lead times. International Journal of Physical Distribution & Logistics Management, vol.37, pp. 115–130 (2007)
6. Nahmias, S.: Análisis de la producción y las operaciones, México: McGraw Hill (2007)
7. Pratap, M., Singh, H.: Evaluation and Economic Selection of Raw Materials Inventory Control Policies: A Case Study. Journal of Operations Management, 7(4), pp. 21–30 (2008)
8. Priyan, S., Uthayakumar, R.: Continuous review inventory model with controllable lead time, lost sales rate and order processing cost when the received quantity is uncertain. Journal of Manufacturing Systems, pp. 23–33 (2015)
9. Quing, L., He, X., Shaohui, Z.: Periodic-review inventory systems with random yield and demand: Bounds and heuristics. IIE Transactions, no. 40, pp. 434–444 (2008)
10. Roach, B.: Origin of the Economic Order Quantity Formula; Transcription or Transformation? Management Decision, vol. 43, pp. 1262–1268 (2005)
11. Sipper, D., Bulfin, R. L.: Planeación y Control de la Producción, México: McGraw Hill (1998)

ISSN 1870-4069

Optimization of Multiple Response Variables Using the Desirability Function and a Bayesian Predictive Distribution

Jorge Limon-Romero[1], Yolanda Baez-Lopez[1], Diego Tlapa[1], Mitzy Pory-Lugo[1], Ivan Rodriguez-Borbon[2]

[1] Universidad Autónoma de Baja California,
Facultad de Ingeniería, Arquitectura y Diseño, Baja California,
Mexico

[2] Universidad Autónoma de Ciudad Juarez, Chihuahua,
Mexico

{jorge.limon, yolanda, diegotlapa, mitzy.pory.lugo}@uabc.edu.mx,
Ivan.rodriguez@uacj.mx

Abstract. This paper proposes the modification of a technique of simultaneous optimization of multiple response variables that works using a Bayesian predictive distribution, to incorporate different weights to the response variables according to their importance in the products. To achieve this, the desirability function has been incorporated to the original proposal. This research shows by representing different scenarios in one case study taken from literature, that the highest desirabilities and in turn the proposed optimum values in the process operating conditions always moved toward regions where the response variables with the highest weights had the best results, at the expense of performance in variables with the lowest weights.

1 Introduction

When working with product optimization, they are likely to have several quality characteristics that must be considered in the analysis. An alternative is to optimize separately each one of these characteristics, which according to various authors is not recommended because this approach may propose as many different combinations of control variables as response variables are considered in the study, which could generate a conflict to decide between them. The best way to address this situation is to consider all responses simultaneously, for which various techniques have been proposed, some of them more complete than others in relation to the various considerations that must be addressed during optimization. It is possible to find alternatives with different levels of technical complexity, with some of them even included within commercial statistical software, which helps them to be the best known and therefore used, although not necessarily the best options. This paper presents a Bayesian approach proposed by [1] which considers the correlation between the responses of interest and the uncertainty in the estimation of the model parameters in a

formal manner. Furthermore it is considered what was proposed by [2], to incorporate noise variables to the analysis. The contribution that will be made have a direct impact on optimizing multiple responses, since this is a modification that allows incorporate a different weights for each response variable, according to the importance they have on the functionality or cost of the product. This will be accomplished by replacing the probability of conformance as variable to optimize in the original approach by the overall desirability (OD) for every process operating condition being analyzed.

1.1 Background

As already mentioned, several options have emerged to work with the problem of simultaneous optimization of multiple responses, the most common based on the loss function [3][4] on process capability indexes [5][6], on the desirability function [7] [8] on the overlapping of graphical response surfaces and have recently emerged techniques that optimize the probability of concordance [9][1]. Some of these techniques emerged to overcome a weakness that had other options or even to complement its earlier versions evolving on the same idea and incorporating some changes that would make them more complete.

Some aspects to consider by the various alternatives proposed to carry out the simultaneous optimization of the response variables are the correlation between the response variables, they must also consider noise factors that affect the quality characteristics and propose operating conditions of the processes that are robust to them and also must incorporate uncertainty in model parameters, all of this without losing the simplicity. So the right alternative must then be selected considering some of these elements.

There have been many efforts by various researchers to design a tool that includes as many of the above mentioned aspects as possible and this information may be found on different research journals. This study proposes to modify an existing technique using the method proposed by Peterson [1] based on the Bayesian approach, incorporating the robust approach proposed by [2] and incorporating the possibility of adding weights to the response variables that presents the methodology based on the desirability function, thus obtaining a hybrid methodology that considers the Bayesian methodology and desirability function to take the benefits of both proposals, for the simultaneous optimization of multiple responses.

1.2 Desirability Function

The desirability function according to [10] is a technique used for optimization of multiple responses in the analysis of experiments in which multiple responses must be optimized simultaneously. According to [11], the desirability optimization methodology is based on the idea that the quality of a product or process having multiple features, when one of these is outside certain "desired" limits, it is completely unacceptable. The method tries to find process operating conditions that provide the "most desirable" response. The optimal values of the factors are determined from the maximization of the function. A high value of D, which varies between zero and one, indicates the best combination of factors to optimize the system studied. [6] propose using desirability functions which converts the problem of multiple responses into a

problem of a single response; that is, the response analyzed is the overall desirability $D=(d_1(Y_1) \ . \ . \ . \ d_m(Y_m))^{1/m}$ where $(Y_1,...,Y_m)$ are the m responses and $d_1,...,d_m$ are the individual desirabilities. To convert the response d_i to y_i these authors propose the following transformation:

$$d_i = \begin{cases} \left[\dfrac{\hat{y}_i(x)-EI_i}{T_i-EI_i}\right]^s & EI_i \le \hat{y}_i(x) \le T_i, \\ \left[\dfrac{\hat{y}_i(x)-ES_i}{T_i-ES_i}\right]^t & T_i \le \hat{y}_i(x) \le ES_i, \\ 0 \qquad \hat{y}_i(x) < EI_i \quad o \quad \hat{y}_i(x) > ES_i, \end{cases} \tag{1}$$

where s and t are used to choose the desired shape of the transformation and thus reflect the wishes of the experimenter: if large values are taken (s, t ≥ 10) means that the desirability d_i only takes large values when it falls near its target value; small values for s and t (s, t ≤ 0.1) means that any value $\hat{y}_i$ within the range [LS_i, US_i] is also desirable. This can be seen in Figure 1.

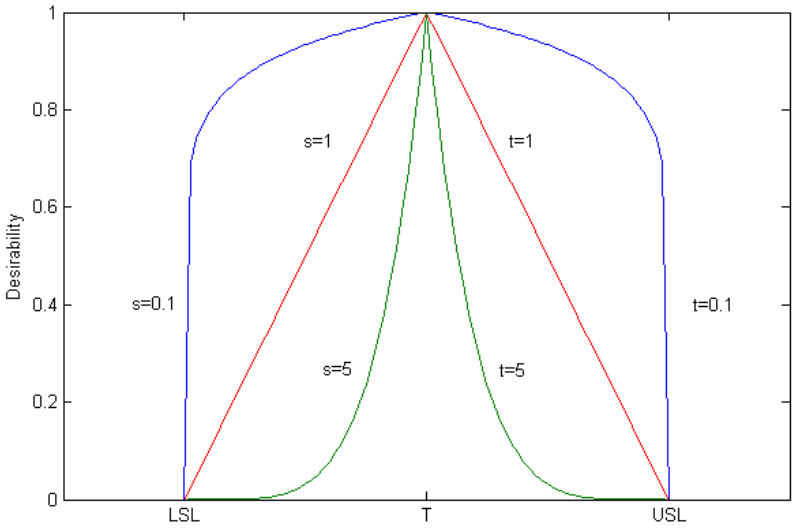


Fig. 1. Desirability function according to different values of s and t.

According to [12] the default value of these exponents is one, suggesting linear increase desirability to the target value (see Figure 1). Also if a response variable has specification on only one side, what must be done is to take the target value (T) equal to the value at which it is considered that no additional gain on quality of the response

is achieved, and in the equation 1 one restriction disappears and figure 1, reduces to one side of the target value (T). According to [15] once the "n" variables (levels of factors and responses) are transformed into values of desirability, these are combined into a single function called Overall Desirability (OD) to find the best set of responses using the following equation:

$$\mathrm{OD} = \left(d_1^{r_1} \times\ d_2^{r_2} \times \ldots \times\ d_n^{r_n}\right)^{\frac{1}{\sum n}} = \left[\prod_{i=1}^{\mathrm{n}} d_i^{r_i}\right]^{\frac{1}{\sum n}}, \quad \textbf{(2)}$$

where r_i refers to weights that represent the importance of each variable with respect to others.

To reach an OD different to zero, it is necessary that all the variables that are being optimized simultaneously have a desirability value greater than zero. On the other hand, if in any operating condition of the process, one of the responses is completely undesirable with $d_i(\hat{y}_i) = 0$, *OD* is also equal to zero, no matter how good the performance is in the remaining variables.

The optimization of multiple responses through desirability function according to [14] has two disadvantages: first, *OD* can be difficult to model because it is a complex function of *m* responses; second, it is difficult to say what the difference between expected values of *D* means except that the bigger is the better. Another disadvantage of this approach according to [15] is that it does not consider the structure of variance-covariance matrix of responses and as it was previously mentioned ignore this information can lead to a no real solution, if indeed the responses have significantly different variances or if they are highly correlated. Other approaches have been proposed based on the desirability function to correct some of these details from the original proposal, see for example, [16, 7, 8, 17, 18].

1.2 The Bayesian Approach

According to [2] the optimization of multiple responses problem is to choose the values of *k* controllable factors x_i, such that $\boldsymbol{Y}$ has certain desired properties. Often the case is presented in which these properties are the vector of responses $\boldsymbol{Y}$ meet specifications. Then let $\boldsymbol{A}$ represents the region of space defined by these specifications responses. Region $\boldsymbol{A}$ may have an arbitrary shape. Then optimizing multiple responses can be summarized in a simple goal, which is to maximize the probability of having the vector of responses within the specification region $\boldsymbol{A}$, for example, maximize the following probability of conformance:

$$\begin{aligned} &\max\ p(x) = p(y \in A \mid x, data) \\ &subjet \quad to: x \in R, \end{aligned} \quad (3)$$

where *R* is the region where the model is valid, which is usually taken as the experimental region defined by the design matrix $\boldsymbol{X}$.

According to [1] based on the typical multivariate regression model:

$$Y = Bx + e, \quad (4)$$

where $\boldsymbol{B}$ is a *pxq* matrix of regression coefficients and $\boldsymbol{x}$ is the qx1 vector of values $\boldsymbol{x}$ in which equation (4) is evaluated. In addition, the vector $\boldsymbol{Y}$ has a multivariate normal

distribution with mean vector **0** and variance-covariance matrix Σ. To consider the uncertainty in model parameters B y Σ it can be used the posterior predictive density $f(y|x,data)$. Using the classical noninformative prior joint distribution for B y Σ and the model of equation (3), the bayesian predictive density for $\boldsymbol{Y}$ can be obtained in closed form. The bayesian predictive density for vector values $\boldsymbol{x}$ where the function is evaluated, is given by a multivariate *t* distribution with ν degrees of freedom (df), as follows:

$$f(y|x,data)=\frac{\Gamma(\frac{\nu+p}{2})}{(\pi\nu)^{p/2}\Gamma(\frac{\nu}{2})}\sqrt{|H|}\{1+\frac{1}{\nu}(y-\hat{B}'x)'H(y-\hat{B}'x)\}^{-\frac{\nu+p}{2}}, \tag{5}$$

where

$$H=(\frac{\nu}{n-q})\frac{\sum^{-1}}{1+x'(X'X)^{-1}x}, \qquad p(x)=\frac{Values\,in\quad A}{Values\quad generated}. \tag{6}$$

Here p is the number of responses to optimize, $\boldsymbol{X}$ is the qxn matrix formed by the n(x_i) vectors of covariates, $\hat{B}$ is the least squares estimation of B, $\Gamma(.)$ refers to the gamma function, $\hat{\Sigma}$ is the usual estimator of Σ calculated using the residuals of multivariate regression adjustment, $\nu=n-p-q+1$ and n is the sample size.

According to [1] because equation 5, follows a multivariate t-distribution, is easy to simulate the $\boldsymbol{Y}$ values from this predictive density. [19] in his book Multivariate statistical simulation mentions that it is possible to simulate a random variable t - varied $\boldsymbol{Y}$, by simulating a multivariate normal random variable and an independent Chi-square random variable [1] If $\boldsymbol{W}$ is a normal random variable with mean vector equal to zero and variance-covariance matrix equal to H^{-1} and additionally $\boldsymbol{U}$ is a Chi-square random variable which is independent of $\boldsymbol{W}$, then:

$$Y_j=(\sqrt{\nu}W_j/\sqrt{U})+\mu_j \qquad \text{for } j=1,\ldots,p, \tag{7}$$

where $\boldsymbol{Y_j}$ is the jth element of $\boldsymbol{Y}$, W_j is the j-th element of $\boldsymbol{W}$ and $\hat{\mu}_j$ is the jth element from $\hat{\mu}=\bar{B}z(x)$. From the above mentioned by [1] it follows that $\boldsymbol{Y}$ follows a multivariate t distribution with ν degrees of freedom.

2 Case Study

To illustrate the operation of the methodology, the experimental data proposed by [20] are taken, which considers three design variables, x_1: reaction time, x_2: temperature and x_3: percentage of catalyst and two response variables, y_1: conversion rate and y_2:

thermal activity using a central composite design with six central runs. In this case x_1 was considered as noise variable (x_n), so the optimization will be performed in relation to the control variables (x_c) x_2 and x_3. Data are shown in Table 1.

Table 1. Experimental results.

Control Variables			Response Variables	
Reaction time	Temperature	% Catalyst	Conversion rate	Thermal Activity
X_n	X_2	X_3	Y_1	Y_2
-1	-1	-1	74	53.2
1	-1	-1	51	62.9
-1	1	-1	88	53.4
1	1	-1	70	62.6
-1	-1	1	71	57.3
1	-1	1	90	67.9
-1	1	1	66	59.8
1	1	1	97	67.8
-1.682	0	0	76	59.1
1.682	0	0	79	65.9
0	-1.682	0	85	60
0	1.682	0	97	60.7
0	0	-1.682	55	57.4
0	0	1.682	81	63.2
0	0	0	81	59.2
0	0	0	75	60.4
0	0	0	76	59.1
0	0	0	83	60.6
0	0	0	80	60.8
0	0	0	91	58.9

Based on the information from the previous experiment the magnitude of subdivisions in the control factors to be used in the simulation to create a fine grid to generate response vectors using equation (7) is determined. These response vectors were used to calculate the corresponding desirabilities. In this case study we worked with a 11^2 array, which means that 121 different combinations of control factors x_2 and x_3 are used with the following levels for each variable:

{-1.682, -1.3456, -1.0092, -0.672, -0.3364, 0, 0.3364, 0.6728, 1.0092, 1.3456, 1.682}.

Once the number of combinations to be considered in the simulation is fixed, it is required to determine the number of runs to be simulated in each combination using equation (7) taking the adjusted regression models based on the original design. This time 10,000 vectors of two responses (y_1, y_2) were generated in each one of the 121 combinations.

At this point it is also addressed the robustness of the process, considering that for each of the combinations of x_2 and x_3, while they remain fixed in the simulation of the 10,000 response vectors, the variable x_1 takes a different value in each iteration, which is generated from a normal distribution (0, 0.1), whereby a vector of covariates x_1, x_2

and x_3 is formed. With these 10,000 response vectors the desirabilities are calculated using equation 1 in each of the combinations of the control variables, comparing the values of the response variables against the corresponding specifications which are presented in Table 2.

Table 2. Specifications of the response variables.

Response	LSL	Target	USL	Kind of variable
y_1	80	110	-	The bigger the better
y_2	50	57.5	65	The nominal is best

Once the individual desirabilities have been calculated the dimension of the problem is reduced by calculating *OD* for each vector generated in each of the 121 combinations using Equation 2. In this case study four scenarios were proposed by changing the levels of importance of the response variables. In scenario 1 (OD_1) the response variables are considered equally important, while in scenarios 2 (OD_2), 3 (OD_3) and 4 (OD_4) is fictitiously considered that response variables have different importance, which is reflected in different weights in order to assess how the optimal point reacts to these weights. Table 3 shows the weights used for each response variable in each one of the proposed scenarios.

Using the *OD* values, polynomials regression models were adjusted using Minitab 16 to represent the relationship between them and the control variables for the four scenarios under analysis. As an example, equation 8 shows the adjusted model for OD_3.

The adjusted R^2 for all four models were greater than 96% and for this reason the optimization (maximization) of these regression models was carried out. In this case, the optimization was realized with the help of Excel Solver using the Generalized Reduced Gradient method (GRG) to determine the best levels of the control variables with the maximum *OD* for every one polynomial model. To initiate the search 15 values for control variables were used. The proposed optimal process operating condition and the *OD* achieved in each model are shown in Table 4.

Table 3. Weights considered for the response variables in each scenario.

	Weight	
Scenario	Y_1	Y_2
OD_1	0.5	0.5
OD_2	0.6	0.4
OD_3	0.8	0.2
OD_4	0.2	0.8

$$MOD_3 = 0.155 + 0.0708\,x_2 + 0.0764\,x_3 - 0.118\,x_2x_3 + 0.320\,{x_2}^2 - 0.0971\,{x_3}^2 + 0.0144\,{x_3}^3x_2 - 0.0299\,{x_2}^4 + 0.0208\,{x_3}^4 - 0.00812\,{x_3}^4{x_2}^2, \quad (8)$$

$R^2 = 96.9\%$.

Table 4. Desirability Values in every optimal process operating conditions.

Scenario	Weights		Optimal values		*OD*
	Y_1	Y_2	X_2	X_3	
OD_1	0.5	0.5	1.682	-0.6055	0.8022
OD_2	0.6	0.4	1.682	-0.5710	0.8513
OD_3	0.8	0.2	1.682	-0.5208	0.9737
OD_4	0.2	0.8	1.5073	-0.7756	0.7140

Finally, optimal proposed process operating conditions are tested using a matlab program very similar to that used in the initial simulation with a small change to allow that the control variables had fixed values, which corresponded to the optimum values proposed by the Solver; to do this 10,000 iterations were used. The simulation was repeated 30 times in order to know how the respective individual desirabilities for every response variable changed. With these values graphical comparisons of the two response variables for the four scenarios were made to see if there was a change on their desirabilities due to the different weights they had in each scenario. Figure 2 shows the behavior of the desirability associated to Y_1 (D_1) in the four scenarios analyzed. The graph shows that generally D_1 has the highest values in scenario OD_3, which is caused because this response variable had a weight of 0.8, while in the other scenarios Y_1 had lower weights. The opposite occurs in OD_4 because is in this scenario where the lowest weight was given to this response variable, reflected in turn in the lowest values of D_1.

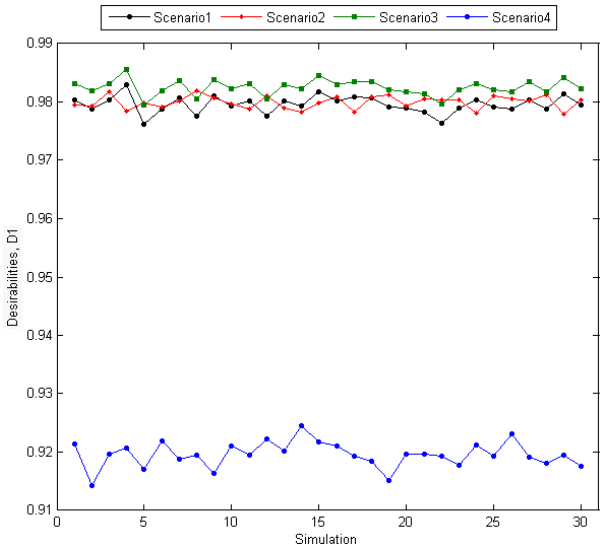


Fig. 2. Comparison of D_1 values in the four scenarios.

In the case of behavior of desirability associated to the variable Y_2 (D_2), the four scenarios can be seen in Figure 3. It is noted that in D_2 the highest desirabilities were in the scenario OD_4 in the 30 repetitions of validation, this is because in this scenario

the highest weight (0.8) was given to this variable. In general, similar conclusions than those presented in figure 2 can be made with D_2 going up or down depending on the weights given to Y_2 in every scenario.

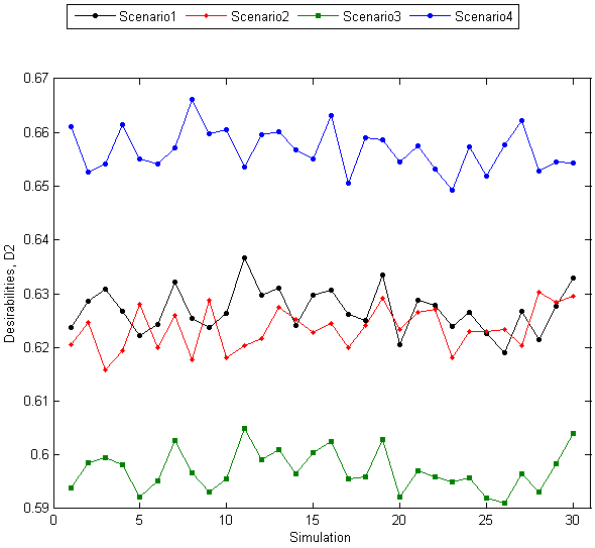


Fig. 3. Comparison of D_2 values in the four scenarios.

With the previous figures the hypothesis that the best results achieved in the different proposed scenarios tend to favor the response variables with highest weights is confirmed graphically.

Table 5. Confidence intervals for the desirabilities in each response variable.

Scenario	*Weigths*		*Average*		*Confidence interval (95%)*	
	Y_1	Y_2	D_1	D_2	D_1	D_2
OD_1	(0.5)	(0.5)	0.9795	0.6269	(0.978972, 0.980101)	(0.625400, 0.628480)
OD_2	(0.6)	(0.4)	0.9799	0.6236	(0.979508, 0.980352)	(0.622063, 0.625037)
OD_3	(0.8)	(0.2)	0.9825	0.5972	(0.981972, 0.982994)	(0.595745, 0.598568)
OD_4	(0.2)	(0.8)	0.9196	0.6568	(0.918730, 0.920397)	(0.655318, 0.658302)

Table 5 shows the 95% confidence interval estimates of individual desirabilities, based on the 30 values obtained during validation. With these intervals it can be seen that there is a statistically significant difference in individual desirabilities of the response variables in the four scenarios presented. In general, the limits of the intervals are shifted up as the weights are increased and on the contrary, the limits decreased when the weights also decreased. Thus, according to these results the proposed

methodology reacted according to the weights given to each response variable, searching the process operating conditions, where the variables with the highest weights behave better.

3 Conclusions

According to the results achieved in this and another case study not reported in this study it is possible to conclude that the methodology proposed by [1] and supplemented by [2] works well with the proposed modification to consider that the response variables have different weights, which in turn can modify the optimum results achieved, searching process configurations with better behavior in the variables with the highest weights. For this reason it makes possible to analyze scenarios where response variables are not equally important, so that the original methodology becomes more flexible and attractive in many practical cases, since in its original version is not designed to incorporate these considerations. Other advantage is that with the proposed modification to incorporate the desirability function, a problem can be solved using different combinations of weights on the response variables, which would provide to engineers and researchers different solutions with respect to the proposed optimal process operation condition, which can enrich decision making.

References

1. Peterson, J. J.: A Posterior Predictive Approach to Multiple Response Surface Optimization. Journal of Quality Technology, pp. 139–153 (2004)
2. Miro-Quesada, G., Del Castillo, E., Peterson, J. J.: A Bayesian Approach for Multiple Response Surface Optimization in the Presence of Noise Variables. Journal of Applied Statistics, pp. 251–270 (2004)
3. Ames, A. E., Matucci, N., Macdonald S., Szonyi, G., Hawkins, D. M.: Quality Loss Function for Optimization Across Multiple Response Surfaces. Journal of Quality Technology, pp. 339–346 (1997)
4. Ch'ng, C. K., Quah S. H., Low H. C.: Index C*pm in multiple response optimization. Quality Engineering, pp. 165–171 (2005)
5. Plante, R. D.: Process capability: a criterion for optimizing multiple response product and process design. IIE Transactions, pp. 497–509 (2001)
6. Derringer, G., Suich, R.: Simultaneous optimization of several response variables. Journal of Quality Technology, pp. 214–219 (1980)
7. Ortiz, F., Simpson, J. R., Pignatiello, J. J., Heredia-Lagner.: A Genetic Algorithm Approach to Multiple-Response Optimization. Journal of Quality Technology, pp. 432–450 (2004)
8. Lee, M. S., Kim, K. J.: Expected Desirability Function: Consideration of Both Location and Dispersion Effects in Desirability Function Approach. Quality Technology and Quantitative Management, pp. 365–377 (2007)
9. Chio, C.H., Hamada, M.: Analyzing Experiments with correlation multiple responses. Journal of Quality Technology, pp. 451–465 (2001)

10. Bahloul, B. Ali L., M., Sfar, S.: A novel Approach for the development and optimization of of self emulsifying drug delivery system using HLB and response surface methodology: Application to fenofibrate encapsulation. International Journal of Pharmaceutics, pp. 341–348 (2014)
11. De Figueiredo, A.K., Rodríguez, L.M., Riccobene, I.C., Nolasco, S.M.: Analysis of the Performance of a Dehulling System for Confectionary Sunflower Seeds. Food and Nutrition Sciencies, pp. 541–548 (2014)
12. Gutierrez, H., De la Vara, R.: Análisis y Diseño de Experimentos. México, Mc Graw Hill (2012)
13. Chiao, C. H., Hamada, M.: Analyzing Experiments With Correlated Multiple Responses. Journal of Quality Technology, pp. 451–465 (2001)
14. Vera, C. L., M. De Zan, M., Cámara, M. S., Goicoechea, H. C.: Experimental design and multiple response optimization. Using the desirability function in analytical methods development. Journal Talanta, pp. 123–138 (2014)
15. Ko, Y. H., Kim, K. J., Jun, C. H.: A new loss function-based method for multiresponse optimization. Journal of Quality Technology, pp.50–59 (2005)
16. Del Castillo, E., Montgomery, D.C., McCarville, D.R.: Of self emulsifying drug delivery system using HLB and response surface methodology: Application to fenofibrate encapsulation. International Journal of Pharmaceutics, pp. 341–348 (1996)
17. Simsek, B., Tansel, Y., Simsek, E. H.: A fullfactorial desing based desirability function approach for optimization of properties of C 40/50 concrete class. Journal of Mathematical and Computational Applications, pp. 330–339 (2013)
18. Zhang, X., Lu, X., Li, S., Zhong, M., Shi, X., Luo, G., Ding, L.: Investigation of 2,4-dichlorophenoxyacetic acid adsorption onto MIEX resin: Optimization using response surface methodology. Journal of the Taiwan Institute of Chemical Engineer, pp. 1835–1841 (2004)
19. Jhonson, M. E.: Multivariate Statistical Simulation: A Guide to Selecting and Generating Continuous Multivariate Distributions. USA: John Wiley & Sons, Inc, (1987)
20. Myers, R. H., Montgomery, D. C.: Response Surface Methodology Process and Product Optimization Using Designed Experiments. New York: Wiley lnterscience (1995)

ISSN 1870-4069

Arabic Text Detection in News Video Based on Line Segment Detector

Sadek Mansouri, Mbarek Charhad, Mounir Zrigui

LATICE Laboratry, Tunisia

mansouri_sadek@hotmail.fr, mbarek.charhad@gmail.com,
mounir.zrigui@fsm.rnu.tn

Abstract. Text embedded in video sequences is very important to semantic indexing and content-based retrieval system, especially for large scale news collection. However, its detection and extraction is still an open problem due to the variety of its size and the complexity of the backgrounds. In this paper, we propose an approach for automatic Arabic-text localization based on a novel method for text-line detection. On the first stage, we use a line segment detector to detect candidate text lines. Then, we propose a word segment identification algorithm based on specific features for Arabic text in order to remove non-text lines. The last stage concerns the text line estimation and text detection in video frames. Experiment results, that we drove on a large collection of video images issued from news broadcasts show the excellent performance of our approach for text detection with different character sizes, directions and styles even in case of complex image background.

Keywords: Arabic text detection, line segment detection, baseline estimation.

1 Introduction

With the development of a big Arabic news channels, News video archives keep increasing in size every day and require more efficient tools for indexing and searching to facilitate access to these collections. Textual patterns embedded in video frames provide high-level information that seems to be a good way for semantic video annotation. However, its detection is still an open problem. This difficulty comes from the variation of style and size and complexity background.

Many methods for text detection and localization have been proposed during the last few years based on different architectures, feature sets, and studies characteristics. These can generally be classified into three categories: connected component-based, edge-based, and texture-based.

The first category assumes that the text regions have a uniform color. In the first step, these methods perform in a color reduction and segmentation in some selected color channel as the red channel in or in color space as Lab space. Then they calculate the similarity of different color values to group neighboring

pixels of similar colors into text region. Shivakumara et al. [1] extract connected components (CCs) using K-means clustering in the Fourier-Laplacian domain, and remove false detections using edge density, text straightness and proximity. Zhuge et al. [2] present a CC based algorithm which applied Maximally Stable Extremal Regions (MSER) as basic character candidates. Text CCs are then grouped into text regions using same geometric rules, and non-text regions are excluded based on corner detection and multi-frame verification.

The edge-based methods use some characteristics of text such as contrast of edge between texts, the background and the density in stroke to detect the boundaries of the candidate text regions. Then, non-text regions are removed by text verification process including some heuristic rules and geometric constraints.

The authors in [3] applied a Sobel filter to detect contour on video frames. Hence, they use the morphological operations to connect the edges together. Thereafter, the candidate's regions that respect the geometric constraints are selected to obtain the coordinates of the text boxes.

The method proposed by Yang et al. in [4] employ an edge based multi-scale text detector to extract text candidates that are then refined using an image entropy-based filter. Support Vector Machine (SVM) is applied as verification procedure to remove false alarms.

The texture-based methods take into account the fact that text regions have special texture features different from other object of background. The first stage is to extract texture pattern of each block in image by applying Fast Fourier Transform, Discrete Cosine transform, wavelet decomposition, and Gabor filter. Then a classification process is applied using k-means clustering, neural network and SVM in order to group each block into text and non-text region.

In [5] the authors propose a method using multi-oriented text detection which is based on the discontinuity of the text regions. To do this, they applied a Sobel mask and a Laplacian filter. Thereafter, Bayesian classifier is used to classify candidate pixels into text and non text regions. These methods face difficulties when the text is embedded in complex background or touches other objects which have similar structural texture to texts. Compared to Latin and Chinese text, few attempts have yet been designed to detect embedded text in Arabic news videos.

In this paper, we propose a novel approach for automatic Arabic text detection in news videos frames using a specific geometric feature of Arabic text called baseline in order to perform detection task. The baseline is defined as the imaginary line which connects all the characters of a word as shown in fig.1. Major contributions for baseline estimation have already been proposed in the field of printed and handwritten document. To the best of our knowledge, our approach is the first which use baseline for Arabic text detection in news video.

The remainder of this paper is structured as follows: In section 2, we discuss works related to text detection and localization. Section 3 presents our proposed approach and its different stages. Section 4 exposes experiment results and section 5states the conclusion.

Fig. 1. Geometric features of Arabic text.

2 Related Works

Unlike Latin and English text, existing methods designed to detect and extract the Arabic text are very few, some approaches have been proposed during the last years.

Ben Halima and al. [6] propose a hybrid approach which combines color and edge to detect Arabic text. Firstly, a multi-frame integration method is applied in order to minimize the variation of the background of the image. Second, a set feature of color and edge is used to localize the text areas.

Alqutami et al. [7] use Laplacien operator to find an edge and k means algorithm to classify all pixels into text or not text region. For regions text, they apply a projection profile analysis to determine the boundary of text block. A similar approach was also presented by Moradi et al [8], a Sobel operator is used to extract edge. Then Morphological dilation is performed to connect the edges into clusters Finally, A histogram analysis is examined to filter text areas.

Sonia et al. [9] propose three methods for Arabic text detection based on machine learning algorithms. First, A Convolution Neural Network is employed for extracting appropriate text, image features and clustering text and non-text images. The two other proposed methods are based on multi-exit boosting cascade. They learn to distinguish text and non-text areas using Multi-Block Local Binary Patterns (MBLBP) and Haar-like features. The experimental results show that the neural network-based method outperforms the other proposed methods.

Recently, Oussama et al. [10] use SWT operator to extract connected component (CC) text candidate. The CCs are grouped based on heuristic rules. Then Convolution auto-encoders and SVM classifier are applied in order to remove non text regions.

Our proposed approach differs from these approaches by employing a specific signature of Arabic text called baseline for embedded text detection in Arabic news videos frames.

3 Proposed System

The whole procedure of our text detection method is mainly divided into three stages as shown in fig.2. The first stage focuses on segmentation process based temporal information of video content. The second stage, presents the main contribution of our proposed approach including three steps: line segment detection, words segment selection, baseline estimation and text localization. In first step text lines and non-text lines in a video frame are identified based on line segment detection [11] algorithm. Second, we apply same heuristic rules based on geometric proprieties of Arabic words in order to remove false detections of text lines. Then, a linear regression method is used to estimate baseline and localize text region. Refinement stage aims to remove false detections.

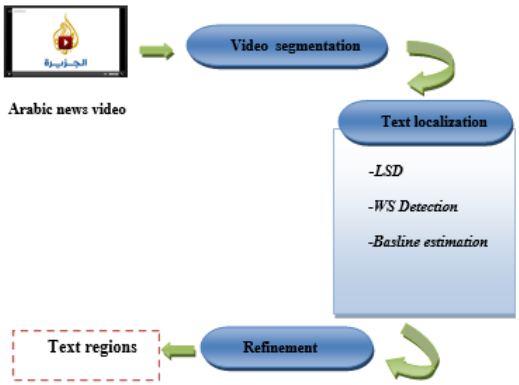


Fig. 2. Global overview of the proposed approach.

3.1 Video Segmentation

To analyze and understand its contents, the video needs to be parsed into segments. Most existing video database systems start with temporal segmentation of video into a hierarchical model of frames, shots and scenes. In this work, we applied a temporal segmentation based on the following assumption the text in the image requires at least two seconds being readable by the user, to generate shots. Then for each video shot, the middle image is selected as a key-frame.

3.2 Text Localization

Starting with the fact that the Arabic text is cursive, we define baseline as a set of word segments described by lines segments. The first step consists of detecting line segment in each key-frame.

Line Segment Detection To do this, we use the line segment detection algorithm that is proposed by Grompone Von Gioi [12]. This algorithm defines a line segment as a region called line-support region and it is based on the information of the gradient direction. Starting with the gradient image, we take a pixel which has a higher gradient magnitude as seed point. Then each adjacent pixel (Ap) which verifies the condition (1) will be added into line-support region and region angle is updated as formula:

$$abs(ang(Ap)le - \theta region) < \tau, \tag{1}$$

$$\theta region = arctan(\frac{\sum_i sin(ang_i)}{\sum_i cos(ang_i)}). \tag{2}$$

This step is repeated until no new pixel can be added to the line-support region. Then, the rectangular approximation of every line-support regions was extracted in order to determine the line segment. Each rectangle was defined by its center, length, width and orientation. In Line segment detection algorithm, the centroid of mass of the rectangular approximate is chosen as the center and the first inertia axis is used to determine the orientation of rectangle. The length and width are chosen in the way that covers the line-support region. Finally, to validate each line segment, a confidence index is calculated based on a contrario model that is proposed by [13].

Word Segment Detection Line segment detection algorithm provides excellent results for line segment detection as shown in Fig2 (b). However, the obtained line segments are likely to be fragmented and touch other non-text objects. To solve this problem, we propose to use some heuristic rules as follows:

Rule 1 Let consider N as the set of detected Lines segments in the image, a line segment where (j N) is considered a candidate word segment if it meets the following conditions:

$$\Theta < \Delta\Theta, \tag{3}$$

$$L < \Delta l, \tag{4}$$

where $\Delta\Theta$ threshold over the direction and Δl the maximal length of the segment that we should detect.

Rule 2 it is difficult to determine which line segment is a word segment based only on the length and orientation. More detailed information is required, including the relationship between line segments. To this end, we define two types of distance: horizontal distance between adjacent line segment (hd) and vertical distance between parallel line segments (vd). Each distance will be used to remove false detection of word segments. Since words segments have been successfully extracted, the last step consists to estimate the baseline using on linear regression method.

Estimating the baseline is a useful task for the reader as well as for Arabic text extraction and recognition. However, its detection is a challenging task due to the wide variety of text visibility, such as variations in font and style and different lighting conditions. Major contributions have already been proposed in the field of printed and handwritten document [14]. The vertical projection is a common method [15] which based on the fact that the word was horizontally aligned and separated by a similar distance between them. Consequently, the baseline is determined according to the maximal peak in the histogram of pixels. All the pervious approaches work with binary image and it cant automatically detect the baseline in video frames that have several challenges such as condition acquisition and complexity background.

In our approach, the baseline is determined based on linear regression method. Let consider c=$\{c_1, c_2, ..., c_n\}$ where $c_i = (x_i, y_i)$ represents the center of word segments within the same direction. The baseline equation is defined by y=ax+b, where:

$$a = \frac{\Sigma_i y_i(\Sigma_i x_i) - n\Sigma_i x_i y_i}{D}, \tag{5}$$

$$b = \frac{(\Sigma_i x_i y_i)\Sigma_i x_i - (\Sigma_i y_i)\Sigma_i x_i^2}{D}, \tag{6}$$

$$D = (\Sigma_i x_i)^2 - n\Sigma_i y_i^2. \tag{7}$$

According to baseline coordinates, the next step of our approach consists of localizing text regions in video frames and representing each region by rectangular bounding box as shown in Fig 3(d).

3.3 Refinement

At this stage, we design a set of heuristic rules based on statistical and geometric properties of text regions to filter out false detections. First of all, we remove candidates regions with very large and very small aspect ratio. Then, we discard the detected regions which are located at the border of the image such as logo of TV channel. We note that the dynamic text will not be taken in account in this work.

4 Experimental Setup

4.1 Corpus

The proposed approach for Arabic-text detection has been tested on a large collection of video news frames. These videos are selected form four Arabic news channel: Al Arabia, Aljazeera (QATAR), AL WATANIYA (TUNISIA), Al Mayadeen (LIBANON) and characterized by the diversity of text pattern such as font, style, position size and background complexity in order to evaluate the robustness of our approach. Our dataset consists of 4000 frames distributed on two sets:

Fig. 3. Steps of text detection: (a) original image, (b) CC extraction, (c) candidates text regions and (d) final detection.

Dataset 1(high definition) a set of 2000 frames extracted from Aljazeera and Al Arabia channels. These channels provide an image resolution that is substantially higher than that of standard-definition (SD).

Dataset 2(standard definition) a set of 2000 frames collected from Al Wataniya 1 and Al Maydeen channels with low resolution. As the evaluation measures we have used recall, accuracy and false alarm.
As the evaluation measures we have used recall, precision and false alarm.

Table 1. Evaluation for our text detection method.

Dataset	Method	Recall	Precision	False alarm
Dataset 1(HD)	Our approach	0.72	0.81	0.28
	Epshtein [16]	0.6	0.63	0.37
Dataset 2 (SD)	Our approach	0.63	0.76	0.37
	Epshtein [16]	0.5	0.59	0.41

4.2 Results

Table 1 shows the experimental results of our method that we drove on two types for dataset. According to these results, it is clear that the proposed approach achieves good results for text detection using Dataset1 (HD) because these types

of channels provide an excellent quality of graphic text. Moreover, our method outperforms the method proposed in [16].

Fig. 4 shows some examples of text localization in video frames. We note that detected regions text will be filtered according to the minimal number of word segment.

Fig. 4. Steps of text detection: (a) original image, (b) CC extraction, (c) candidates text regions and (d) final detection.

We note that our method face difficulties when other objects have similar geometric characteristic to Arabic text as shown in fig 5.

Fig. 5. Same false detections of our method.

5 Conclusion

In this paper, we have presented a novel approach for text localization in the Arabic news video. A specificity of our proposal is to use the geometric features for Arabic text such as word segments and baseline in order to enhance text detection in video frames. Experimental results have shown that the proposed method is able to detect embedded text with different text appearances and complex backgrounds in HD channels and achieved higher recall and precision score than SD channels.

In future work, we plan to use other visual features to enhance detection task especially for video frames with low resolutions.

References

1. P. Shivakumara, T. Q. Phan, C. L.Tan: New Fourier-Statistical Features in RGB Space for Video Text Detection. IEEE transactions on Circuits and Systems for Video technology (CSV), pp. 1520–1532 (2010)
2. Y. Z. Zhuge, H. C. Lu: Robust video text detection with morphological filtering enhanced MSER. Journal of Computer science and Technology, pp. 353–363 (2015)
3. Poignant, J., Besacier, L., Quenot, G., Thollard, F.: From Text Detection in Videos to Person Identification. In: IEEE International Conference on Multimedia and Expo (ICME) (2012)
4. H. Yang, B. Quehl, H. Sack: A Framework for Improved Video Text Detection and Recognition. International Journal of Multimedia Tools and Applications (MTAP), pp. 217–245 (2012)
5. Shivakumara, P., Sreedhar, R. P., Phan, T. Q., Lu, S., Tan, C. L.: Multioriented Video Scene Text Detection through Bayesian Classification and Boundary Growing. IEEE Transactions on Circuits and Systems for Video Technology (2012)
6. A.M. Alimi, M. Ben Halima, H. Karray, A. Fernández Vila: Nf-savo: Neuro-fuzzy system for Arabic video OCR. International Journal of Advanced Computer Science and Applications, vol. 3, no. 10, pp. 128–136 (2012)
7. A. Alqutami, A.M.A. Ahmad, J. Atoum: A robust algorithm for Arabic video text detection. In: Proceedings of International Congress on Computer Applications and Computational Science, Advances in intelligent and Soft Computing, Bali, Indonesia (2011)
8. S. Zhang, C. Zhu, J. K. O. Sin, P. K. T. Mok: A novel ultrathin elevated channel low-temperature poly-Si TFT. IEEE Electron Device Lett., vol. 20, pp. 569–571 (1999)
9. M. Moradi, S. Mozaffari, A.A. Orouji: Farsi/Arabic text extraction from video images by corner detection. In: Proceedings of 6th Iranian Conference on Machine Vision and Image Processing, Isfahan, Iran (2010)
10. Yousfi, S., Berrani, S. A., Garcia, C.: Deep Learning and Recurrent Connectionist-based Approaches for Arabic Text Recognition in Videos. ICIP (2014)
11. O. Zayene, J. Hennebert, S. M. Touj, R. Ingold, N. E. Ben Amara: Text Detection in Arabic news Video Based on SWT Operator and Convolutional Auto-encoders. In: Proc of 12th IAPR Workshop on Document Analysis Systems (2016)
12. R. Grompone Von Gioi, J. Jakubowicz, J.M. Morel, G. Randall: LSD: a fast line segment detector with a false detection control. IEEE Transactions on Pattern

Analysis and Machine Intelligence, vol. 32, no. 4, Article ID 4731268, pp. 722–732 (2010)
13. A. Desolneux, L. Moisan, J. Morel: Meaningful alignments. International Journal of Computer Vision, vol. 40, no. I, pp. 7–23 (2000)
14. Almuallim, H., Yamaguchi, S.: A method of recognition of Arabic cursive handwriting. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), 9(5):715–722 (1987)
15. Abu-Ain, T., Sheikh Abdullah, S., Bataineh, B., Omar, K., Abu-Ein, A.: A Novel Baseline Detection Method of Handwritten Arabic-Script Documents Based on Sub-Words. Soft Computing Applications and Intelligent Systems, Communications in Computer and Information Science, 67–77 (2013)
16. B. Epshtein, E. Ofek, Y. Wexler: Detecting text in natural scenes with stroke width transform. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2010)

ISSN 1870-4069

Learning-to-Rank for Hybrid User Profiles

Houssem Safi, Maher Jaoua, Lamia Belguith Hadrich

ANLP Research Group, MIRACL Laboratory
Faculty of Economics and management of Sfax,
Tunisia

safi.houssem@gmail.com, {Maher.Jaaoua, la.beguith}@fsegs.rnu.tn

Abstract. In the context of the Personalized Information Retrieval method applied to the Arabic language, this work consists in presenting a personalized ranking method based on a model of supervised learning and its implementation. This method consists of four steps, namely, the user's modeling, the document / query / profile matching, the learning to rank and the result classification. Thus, we proposed a hybrid approach of the user's modeling that relies on both multidimensional and conceptual representations by exploiting Arabic semantic resources. Therefore, to determine the similarity between the document and the profile, we used a learning model that exploits the users' explicit pertinence judgments. In this context, we have proposed learning semantic features related to the user's profile (represented by hierarchies of concepts). The predicted model will then be used in the ordering phase to classify other documents that result from a new query submitted by the user. In this context, we have proposed a novel multi-objective function to order the documents (based on the classic Retrieval Status Value function and the predictive personalized Retrieval Status Value function). Finally, we have explained the evaluation results of the predictive model and the ranking method. These evaluations, which were made based on a training corpus and a test corpus, led to some interesting results. Indeed, the proposed semantic learning criteria connected to the user profile have a significant impact on the performance of our personalized document ranking system.

Keywords: document ranking, learning to rank, hybrid profile, personalized retrieval status value.

1 Introduction

Personalized Information Retrieval (PIR) is one of the best sources of information for acquiring user-based information more precisely and efficiently [1]. PIR is a novel technique where many techniques have been developed and tested; however, many issues and challenges are still to be explored. The most common encountered difficulties, when searching for information, are [2]:

- Problems with the data themselves,

- Problems faced by the users who try to retrieve the data they want,
- Problems in understanding the context of the search queries and
- Problems with identifying the changes in the user's information need.

Moreover, many PIR methods have been discussed in literature [3]. The problems with the existing methods explained in the following observations [3] are the user's protection and the unnecessary disclosure of his profile.

Therefore, the major aim of the researchers who are going to work on this issue will be to completely protect the users and introduce new techniques to prevent unnecessary disclosure of their profiles. We need an innovative approach to create a dynamic user profile based on a submitted query. Furthermore, to our knowledge, very little research has been devoted to personalized information for the Arabic language.

For this reason, the work presented in this paper aims at developing a system for PIR which can be adapted to the Arabic language and provide personalized results based on the user's preferences and interests. This system is dubbed SPIRAL (System for Personalized Information Retrieval applied to Arabic Language). The SPIRAL system uses the reformulated queries (the method adopted the reformulation is proposed by [6]) to reorder the documents retrieved by a search engine while taking into account the user profile. Thus, the implementation and evaluation of personalized learning to rank method and the integration of a hybrid user profile are the subject of this work.

The language targeted by this system at the query and returned documents is the Arabic language. The choice of this language is motivated by the fact that Arabic has not received the same interest as other languages, such as French or English. Similarly, in recent years, we have noticed the emergence of Arabic language resources in the field of automatic language processing. Therefore, the integration of these resources into operational systems dealing with the Arabic language is an additional motivation.

In the second section, we will present a brief overview of the Personalized Information Retrieval (PIR). More precisely, we will briefly explain the learning to rank approaches of documents, then we will present a state of the art of the IR applied to Arabic. In the third section, we will deal in detail with the integration of the user's profile in the proposed method of ranking. In the last section, we will provide a description of the learning to rank system as well as an evaluation of our own corpus.

2 Personalized Information Retrieval

The PIR is a general category of search techniques aiming at providing better research results. The solutions for the PIR can generally be categorized into two types, namely *profile-based* [5] and *click-log-based* [5] methods. The profile-based methods improve the search experience with complicated user-interest models generated from the user's profiling techniques. In the click-log based methods, the authors simply impose a preference to clicked pages in the user's query history. One limitation that

reduces its applicability is that it can only work on repeated queries from the same user.

It is emphasized that this work is in the context of the combination of the profile based and click-log-based methods. Thus, the personalization system needs to use all the information about the user (profile, main interests, preferences, information needs) and his research environment [3]. There are mainly tree types of representations of the user profile: Semantics, Multidimensional and Set. The adaptation to the changes in the interest centers, which describe the users, means the upgrading of the user profile. There are two types of user's needs: long-term and short-term profile.

In what follows, we will give a brief review of the learning to rank approaches and a comparison between the models. In addition, we will describe the IR systems applied to Arabic. Finally, we will identify some limitations of these systems.

2.1 Brief Overview of the Learning Approaches to Document Ranking

During the last decade, many algorithms have been proposed to optimize the re-ranking of the search results. These algorithms are generally divided into three categories: pointwise [6], pairwise [7] and listwise [8]. These approaches differ according, first, to their way of considering the input data of the learning system, second, to the type of the variable or judgment of relevance to predict and, third, to the mathematical modeling of the learning problem.

In the pointwise model, each document xi is considered a separate input of the learning model. The judgment of relevance can be an integer or a real score, an unordered class of relevance (not relevant, relevant) or an ordered class of relevance (level 1 relevance <level 2 relevance <...). The judgment of relevance here is a variable that predicts the value which ranks the documents. When the judgment of relevance is an integer or a real score, the learning problem is generally regarded as a linear regression problem. The relationship between the quantitative variable to be explained and the explanatory variables is assumed to be linear.

In the pairwise model, the pairs of documents (x_i, x_j) are considered as an input to the learning step. Each pair of documents is associated with a judgment of preference $y_{i,j}$ with value $-1, 1$. If $y_{i,j} = 1$, then document x_i, which is favorite to document x_j: should be ranked above x_j in the result list. Preference is denoted $x_i \succ x_j$. On the other hand, if $y_{i,j} = -1$, then document x_j is preferred to x_i document and notes $x_j \succ x_i$. The learning problem here is a classification problem, in the particular case of pairs of instances. Therefore, most of the algorithms of this model use adaptations of existing classifiers.

In the listwise model, a complete and ordained list of documents is considered as an input of the learning step. The algorithms provide as output the ordered list of documents or a list of their relevance scores ([8, 9, 10, 11, 12, 13]). The algorithms are divided into two subcategories within this model: those minimizing an error function defined from an IR measurement as MAP (MAP is the average of the average precision of all the queries [16]) or NDCG (Normalized Discounted

Cumulative Gain is defined from the Discounted Cumulative Gain (DCG) [16]) and those minimizing a loss function not related to the IR measurement.

Historically, the Pointwise and Pairwise models have been the first to be proposed (around the early 2000s) while the first studies treating the Listwise model have appeared only recently. Some other research studies have been proposed to compare the learning approaches for the above ranking. The conclusions drawn show that the model list shows more interesting results than the models in pairs or points [14] and [15]. It should be noted that these results were obtained following the analysis of large number of algorithms and large data sets (3.0 for the collection Letor [14, 15]). In addition, the Listwise model is generally regarded as easier to implement. Therefore, we chose to use the list approach in our learning model.

2.2 Information Retrieval Applied to Arabic

Faced with the IR, the Arabic language has recently been addressed by conventional search engines, but it is absent in the semantic search engines. It is within this context that this work proposes to develop a personalized information retrieval system for the Arabic language. This system illustrates the implementation of the PIR method that we have proposed and which distinguishes three stages, namely the user's modeling, reformulation (specifically expansion) query and scheduling results.

The attention paid to the Arabic language is explained by the fact that this language does not receive the same degree of attention as the other languages such as French or English. Moreover, the Arabic language resources are emerging in the search field of automatic processing of language which gives extra motivation to integrate these resources into operational system processing of the Arabic language.

In the implementation of our PIRS, we will try to incorporate language resources developed for Arabic. This consists in integrating a chain of linguistic analysis which, besides helping resolve the language ambiguities, enriches the concepts of the users' queries and profiles.

To solve the morphological and lexical ambiguity, a lemmatizer is suggested to place a light lemmatization. The use of semantic resources for the enrichment (expansion) of the user's query can be a solution to solve the problem of semantic variations and disambiguate the query terms. Indeed, the semantic resources provide resources in the form of semantic relationships. They can extend the search field of a query, which improves the research results.

The use of semantic resources in an IRS may be considered at several levels:

- Before being sent, the user's query can be enriched by the near judged concepts in semantic resource through the use of relationships, such as generalization / specialization, synonyms ...
- The indexing of documents is made using the concepts of the semantic resource and not the keywords.
- Filtering of documents in a particular field to the user profiles ([17, 18, 19]).

It should be noted that the query expansion is a double-edged sword so that improving the research in this event may be accompanied by an information overload problem.

Indeed, the query reformulation or expanding may generate a significant number of terms when using multiple relationships in a semantic resource.

To address this problem, we propose a second alternative based on the user profile concept to reduce the enriched elements during the expansion, in order to remove the ambiguity of some terms and filter the returned documents. Similarly, we propose a third alternative to improve the accuracy of the IR entitled "personalized learning to rank". This alternative, which is based on a hybrid user profile (multidimensional and conceptual), makes it possible for the user to put the classified documents, which are "relevant" according to his profile, at the top of the list.

To our knowledge, there are no PIR systems for Arabic. Most of the developed research studies in the field of IR in Arabic have been particularly interested in the query reformulation step. These studies use the thesauri dictionary and the language resources to substitute and / or disambiguate the query terms. In the following part of this section, we will quote the main research studies in the context of an IR in Arabic, then, we can group them according to two axes. The first axis includes the work using morphological stemming of the query words, while the second includes the studies that exploit the thesaurus dictionary.

In the first axis, Xu and al. evaluated two research strategies of Arabic documents using the ArabTREC corpus as a test corpus. The authors developed a strategy that uses first indexation based on the roots. This method resulted in a slight improvement of the research results. Likewise, these authors showed that the second strategy that is the use of a thesaurus dictionary, dramatically improves the performance of an Arabic IRS [20].

On the other hand, Bessou and al. adopted the scheme notion as a base to substitute the query words with their lemmas at the level of indexing and search steps [21].

In the second axis, we can mention the work of Hammo and al. that used the Koran as thesaurus for the query reformulation [22]. For their part, [23] used the Arabic WordNet as thesaurus to supply the ontology designed for the legal field.

The work of [24] proposed to assist the user with the reformulation of his query by adding nearby morphological forms of the initial query word forms. This addition is based on a similarity calculation of n-grams between the words of the original query and those saved in a lexicon. To index and search for operations, [24] used the services of the Google search engine.

The work of [25] can be summarized in the use of an external resource (Arabic WordNet or AWN) and a morphological analyzer to be reformulated by expanding the user's query that can improve the recall but not the precision of the IRS. As an extension of this work, [26] used a reformulation based on two external resources, namely ADS (Arabic Dictionary of Synonymy) and AWN.

It should be emphasized that the already mentioned research studies have some limitations. Indeed, some studies ([23] and [22]) used semantic resource or ontology for a specific field. Besides, there is non-use of conceptual relationships ontology in some studies. Finally, there is a lack of studies ([23] and [26]) about the contribution of each semantic relationship used in some Arabic query expansion systems.

According to the conducted overview, we can conclude that the enrichment of queries based on external resources is an interesting path the exploitation of which

can improve the results of the IR. In addition, we noticed that the personalization side is absent in the above studies, which is an additional motivation for this work knowing the performance improvements recorded in other languages. On the other hand, we can emphasize that learning to rank is a technique totally unaffordable by PIR systems for the Arabic language which is another motivation for this work. Indeed, semantic learning features from the user profile and those contained in the semantic resources constitute an original and a promising path that can give good performances in the context of IR in Arabic.

To our knowledge, there is no personalized learning to rank systems dedicated to the Arabic language (that is to say there are not works that integrate the user profile). Likewise, it is worth noting that our contributions of this work in the field of PIR revolve around the following points:

- Modeling of a hybrid user profile that relies on both conceptual and multidimensional representations by exploiting Arabic semantic resources.
- Proposing semantic learning criteria connected to the user profile (represented by concept hierarchies). These criteria have a positive impact on the performance of our PIR system.

3 Proposed Method

The objective of the personalized ranking method is to provide the user with an ordered list of documents in response to a query issued by him. The document ranking is a major theme in the IR. Indeed, several studies have been made to establish the appropriate metrics that help determine the optimal order governing the documents returned by a search engine. The many features that were proposed to develop these ranking metrics are the similarity of documents in relation to the query, their importance and their links [15, 27], etc.

Since the proposed method is based on the user profile, it is quite apparent to integrate the profile in the calculation of its similarity with the documents returned by the search engine. It should be noted that the used queries are reformulated and, therefore, they integrate concepts from the profile. It follows that personalization is given a leading role in the result ordering.

To determine the similarity between the document and the profile, we used a learning model that exploits the users' explicit judgment pertinence. This consists in asking the user to assign a relevance class to document that reflects its significance in relation to his needs. In a second phase, we project these judgments on features related to the documents, the queries and the profile.

This projection helps build a predictive model that discerns the relevant documents meeting the user's profile and query. The predicted model will then be used in the ranking phase to classify other document results of a new query submitted by the user.

In the following part of this section, we will introduce the ranking document method that distinguishes four steps, namely, (1) the user's modeling, (2) the document/query/profile matching, (3) the learning to rank, (4) and the result classification as shown in figure 1.

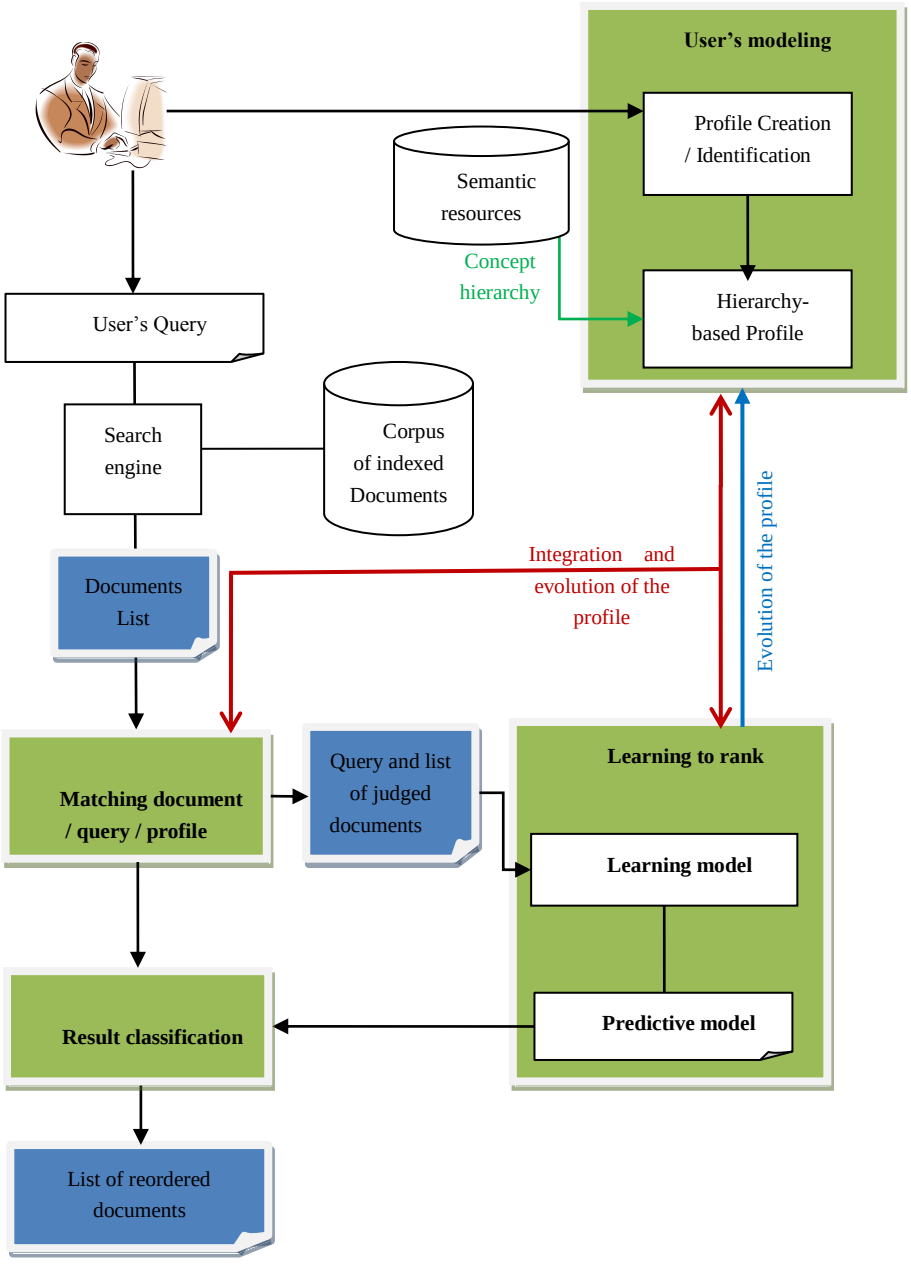


Fig. 1. Personalized learning to rank method

It should be emphasized that, in our ranking document method, we have included the method of document/query/profile matching that was used in [4]. For this reason,

step (2) will be presented in brief while steps (1), (3) and (4) will be described in detail.

3.1 Suggested User's Modeling

In the framework of the proposed ranking method, a user's modeling based on a hybrid representation and built on the user profile is proposed. In this approach, an algorithm which automatically builds a hierarchical user profile is introduced to represent the user's implicit personal interests and domain. It is to represent the domain and the interests with a conceptual network of nodes linked together. This network is made through relationships respecting the linking topology (synonymy, hyponymy and hyperonymy) defined in ontologies (AWN [30] and Amine AWN [31]) and the domain of hierarchies.

It should be noted that our method allows updating the short and long term user profile. The evolution of the user profile in short term is jointly linked to a bounding mechanism of search sessions to examine the change of interest over time. In addition, relevant feedback helps refine the user's preferences and consequently update the short-term profile.

Thus, the capture of changes in the centers of interests is concretized by the addition of the search history (queries and search results that have been appreciated by the user) to the short term profile. Indeed, the proposed method establishes an activation score based on the construction and evolution of a user profile from his judgments of relevance. In this context, each user's query will be added to his profile in the short term. A weight averaging the formula tf * idf, will be assigned to each term derived from the document deemed relevant or very relevant by the user.

Then, the first terms with the largest weight will be inserted in the short-term user profile. The number of these added terms can be determined by an experimentation which achieves the compromise between the size of the centers of interest and its real needs. It should be noted that in this method, an algorithm of the concept score propagation is used to update the weights of the profile concepts.

Indeed, the terms of consulted documents and / or submitted queries are aggregated to the user profile according to a similarity threshold between the document and the user profile. In this phase, we adopt a method which models profile V by R vectors V_i respectively corresponding to R documents d_i judged as relevant by the user. For each new selected document d_i, the V_i dimension, which is the most similar to the profile of document $d_{i'}$ is updated as follows:

$$V_i = V_i + V_{i'};\ V_i = \text{argmax}_{v_i \in v}\ \text{Sim}(V_i, V_{i'}) \text{ with}$$

$$\text{Sim}(V_i, V_{i'}) = \frac{V_i . V_{i'}}{|V_i| . |V_{i'}|}. \quad (1)$$

Only m words $t_v \in V_i$, which have longer weights, are selected for updating dimensionV_i of profile V.

Thus, the long-term user profile enables (implicitly and / or explicitly) to model persistent or recurrent centers of general interests. The evolution process of the long-

term profile is to add or change a context formed by concepts associated with a query sent by the user. Identifying a similar context to the user's profile involves merging them and subsequently updating the long-term profile. A new context is therefore added to the long-term profile if no previously learned context is similar to the context of the query. Likewise, the modification of the long-term profile can be envisaged by enabling the user to explicitly integrate a new domain.

Generally, high levels of hierarchy concepts make it possible to represent the profile in the long term whereas low levels make it possible to represent a high level of specificity of the user profile in the short term.

3.2 Personalized Matching Step

The calculation of the personalized matching score between the document and the profile can be determined by the cosine between both $\vec{D}$ and $\vec{U}$ vectors. At this level, we can set a threshold for RSV (D, U) below which document D will not be retained in the list of results for a given query. This threshold may be determined after a series of experiments to select the documents that best satisfy the user's needs [4].

3.3 The Learning-to-rank Step

The ranking step takes as input a list of documents judged by the user and his profile. The latter is based on a concept hierarchy extracted from the semantic resources. Similarly, the list of documents, which is the training corpus, contains learning features labeled by the user. Thus, the learning phase is based on the optimization of a ranking function that leads to a predictive model.

In what follows, we will describe the learning to rank principle then we will spread out the adopted learning features.

Principle of learning to rank. The classic ranking function is used to classify only the documents that take account of the user's queries in a descending order of relevance. In the case of personalized learning, our contribution is to classify the documents that take account of the queries but also the user profile. Given that our goal is to order a list of documents, the most appropriate model to use in the learning step is the listwise model. This model also has the advantage of evaluating the performance of the algorithms on the basis of IR measurements, as it displays more interesting results than the other models.

The learning to rank is based on two concepts: the representation of the document-query-profile triplet in the feature space and the use of a learning model. The learning to rank process is divided into two phases: a training phase and a testing phase. In the learning phase, the datasets are used by algorithms to automatically learn the ranking functions that serve as models for the prediction of relevance judgments (the chosen scale is three classes of relevance: relevant, slightly relevant and irrelevant). In the test phase, these functions are then used to order the documents returned by the IRS when new queries have been submitted.

The data set used in the learning phase, consists of the query/document/profile triplets. Each triplet (q_i, d_j, u_k) is represented in the feature space by the vector $x_{i,j,k} \in \mathbb{R}^d$ such that $x_{i,j,k} \in \mathbb{R}^d$ and $x_{i,j,k} = [x_{i,j,k}^{(1)} \ldots x_{i,j,k}^{(d)}]$ and associated with a class of relevance $s_{i,j,k}$.

In the test phase, the learned function is used by the ranking system to predict the relevance scores of new triplet query / document / profile which have not been annotated. The ranking model thus returns the relevance of the class for each query / document / profile.

Learning proposed features. The used learning model operates a set of features that depend on the query, the document and the user profile. In order to measure the impact of personalization using the learning technique, we were led to choose learning criteria related to the user profile (represented by hierarchies of concepts). The adopted features can be classified in four categories.

The first category consists in determining the similarity between the query and the returned documents. The features are used to calculate the term frequency (tf) of the original query in the text, the title, the subtitle, the summary, the category and the index of document.

The second category of features includes similar features between enrichment query words and the document. The features help extract the matching frequency of the terms synonyms, the generalization and the specification in the document.

The third category is related to the similarity between the document and the user profile. The purpose of these features is to verify the presence of the short or long term user profile concepts in the text. This feature is based on the tf representing the degree of similarity between the user profile and the document. More precisely we determine the frequency of the centers of interest concepts, of the short and long term profile with the document.

The fourth category includes other contextual features related to documents and query and their statistical characteristics. We can mention, as an example, the number of query words, the number of words in the text, the text length (short, medium or long) as well as the format features (Word, PDF, PowerPoint, etc.).

It should be noted that the learning to rank features consist of one of our contributions in the field of PIR given that, according to our knowledge, there are no research studies that used this type of features.

Relevance class. In the framework of classical IR, the process of judging the information relevance is based on the degree of similarity between the representation of the query and the content of the document found by the system.

However, personalization involves taking into account the user profile as an information source that participates in the judgment of relevance. Thus, relevance can be defined as the adequacy of a document following a given query and a well-defined profile. This notion is subjective because the user's state of knowledge is dynamic. Indeed, for the same user, relevance changes over time while a document can have different types of pertinence for two users who submitted the same query.

To annotate the relevance class of a document, we can borrow the explicit feedback approach of the user. Under this approach, the user directly delivers his interest judgment by giving a relevance value on a graduated scale from the least to the most relevant. In our method, the class of a document compared to a query for a given user can have one of the following words "irrelevant," "medium relevant" or "relevant".

It is noteworthy that we have initially chosen five evaluation degrees, namely "irrelevant," "a little irrelevant," "moderately relevant", "relevant" and "highly relevant". However, we detected two problems of annotation (overlap between the entries) between the first two points "irrelevant" and "moderately relevant" and between the last two "relevant" and "highly relevant". In fact, we found that the users or even experts find it difficult to judge the documents using five rating levels. For this reason, we were led, in a second stage, to keep only three levels.

3.2 The Results Ranking Step

The final result ranking depends on the relevance of the documents in relation to the query and the user profile. This relevance combines two values namely the classic RSV (D, Q) and the predictive personalized RSV (D, Q, U) where D, Q and U are respectively the document, the query and the user profile.

To measure the classic RSV function, we adopt the most known measures from the quantities called tf and IDF. Our choice is justified by the fact that these measures are very successful and very popular in the IR. The weight of a word in a query or in a document is expressed using the tf.IDF measurement. The tf measure is the number of word occurrences within a document, while the IDF measure shows the importance of a word in the considered corpus, such as:

$$IDF(t) = log \frac{N}{n_t}. \quad (2)$$

It is noteworthy that the predictive personalized RSV function is a relevance class which can either be "irrelevant," "medium relevant" or "relevant", whereas the classic RSV function is a score calculated by the cosine function which belongs to the interval [0..1]. Due to the incompatibility of both functions, we have adopted a multi-objective function that promotes first class relevance of the documents. In the case where two documents have the same class relevance, the multi-objective function uses the classic RSV function. Therefore, as a first step, we ranked the documents based on their similarity to the profile. As a second step, we classified the documents with the same relevance class according to their similarities with the query.

4 Implementation and Discussion of the Results

The implementation of the proposed PIR method resulted in three versions. The first version is the query expansion system, the second version, which is a system that integrates the personalized matching module but does not contain the ranking module.

The third version of our system is the "SPIRAL" that includes all the steps of the proposed method. . In this section, we will provide a description of the SPIRAL system as well as an evaluation of our own evaluation corpus.

4.1 Arabic Corpora for Learning and Ranking

Since there are no evaluation standards for personalized access to information, especially for short-term personalization, we proposed context-oriented assessment frameworks based on simulation collections of TREC campaign by simulated users' profiles and search sessions. We have exploited these evaluation frameworks to validate our SPIRAL contribution. For this reason, we have created a large Arabic text corpus entitled WCAT (Wikipedia Corpus for Arabic Text) using the search engine "Lucene"[1]. This corpus is segmented into 30550 text article, extracted from Wikipedia. This corpus contains texts dealing with topics related to the "natural sciences" domain. Moreover, each article has one or more categories related to the root category of "natural sciences". We generated 7200 sub-categories from the "natural sciences" category.

The search engine Lucene is capable of processing large volumes of documents with its power and speed due to indexing. In our system, we used Lucene to index a corpus of documents, analyze the queries, search for the documents and present document results.

In this phase, the indexing step of the corpus consists in stemming words, removing stop words, indexing and extracting key words of each document in the corpus.

We also built our own Arabic Query Corpus entitled "AQC_2", which is composed of 1000 queries submitted by 50 different users and deals with topics related to the "natural science" domain. An Arabic query corpus consists of 90,507 words or 613,021 characters and 3.47 megabyte size. Thus, the evaluation corpus of our system contains different types of queries suggested by various users.

When working on a learning process, it is appropriate to divide an initial corpus into two sub-corpora:

- The learning corpus serves to extract a model or classification from a sufficient occurrence of information;
- The test corpus is used to check the quality of learning from the learning corpus.

In what follows, we will give some features of the learning corpus and the learning evaluation corpus (table 1).

It is emphasized that in the context of evaluating the ranking system, we tested the SPIRAL system for 50 users; each of whom has submitted 20 queries. This gives us a corpus of 1,000 test queries. Therefore, in our assessment of every query, only the first 10 documents returned by the search engine are taken into account, which gives us a test corpus of 20,000 documents.

[1] https://lucene.apache.org/

Table 1. The learning and the evaluation corpora.

	Size of the corpus	Average size of an item	Number of items	Number of Words	Langu age
Learning corpus	65 mega-octets	4 Kilo-octets	20 000	15 333 028	Arabic
Evaluation corpus	35 mega-octets	4 Kilo-octets	10 000	6 159600	Arabic

In what follows, we will present the evaluation results of the SPIRAL system. We used the Weka learning framework to get to know the personalized ranking function of our system that exploits the user profile so as to reorder the documents returned for a given query.

4.2 The Used Indicators of Performance

The indicators of performance are used to evaluate a prediction model; however, the performance of this model can be significantly influenced by the conditions of its experimentation. In this section, we will first describe the different evaluation indicators of the prediction models, then, the standard performance measures. Finally, we will present the cross-validation method that we used to evaluate our learning model.

Standard measures of performance. To evaluate the learning model, we used assessment measures such as the recall, precision and F-measure. In addition, we used the kappa measure which measures the degree of agreement between prediction (predicted classes) and supervision (real classes) after the agreement by chance is removed.

$$\text{Recall}_{(i)} = \frac{\textit{Number of documents assigned correctly to class } i}{\textit{Number of documents belonging to class } i}, \quad (3)$$

$$\text{Precision}_{(i)} = \frac{\textit{Number of documents correctly assigned to class } i}{\textit{Number of documents assigned to class } i}, \quad (4)$$

$$\text{F-Measure}_{(i)} = \frac{2 \text{ x Recall x Precision}}{(\text{Recall} + \text{Precision})}. \quad (5)$$

Cohen's kappa: this coefficient is a statistics which measures the inter-rater agreement for qualitative (categorical) items. It is generally thought to be a more robust measure than the simple percent agreement calculation, since κ takes into account the agreement occurring by chance. The equation for kappa (K) is:

$$\text{Kappa}_{(i)} = \frac{\Theta 1 - \Theta 2}{1 - \Theta 2}, \quad (6)$$

where Θ1 is the relative observed agreement among the raters, and Θ2 is the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. If the raters are in a

complete agreement then $\kappa = 1$. If there is no agreement between the raters other then what would be expected, then, (as given by $\Theta 2$), $\kappa \leq 0$.

It should be noted that the error rate is equal to the difference between the rate of the ideal classification (100%) and the good classification rate:

$$\text{Error Rate} = 100\% - \text{good classification rate.} \quad (7)$$

Cross-validation. Cross-validation, which is sometimes called rotation estimation, is a model validation technique for assessing how the results of a statistical analysis will generalize to an independent data set [32] [28] . It is mainly used in settings where the goal is prediction, and one wants to estimate how accurately a predictive model will perform in practice.

In k-fold cross-validation, the original sample is randomly partitioned into k equal sized sub-samples. Among the k sub-samples, only one is retained as the validation data for testing the model, and the remaining k-1 sub-samples are used as training data. The cross-validation process is then repeated k times (the folds), with each k sub-samples used exactly once as the validation data. The k results from the folds can then be averaged (or otherwise combined) to produce a single estimation. The advantage of this method, over repeated random sub-sampling, is that all the observations are used for both training and validation, and each observation is used for validation exactly once. 10-fold cross-validation is commonly used, [29] but in general, k remains an unfixed parameter.

4.3 Evaluation and Discussion of Learning Model Results

This section focuses on the different experiments carried out for our learning model. Indeed, these experiments are expressed in terms of global accuracy using, on the one hand, the decision trees and, on the other hand, the SVM in addition to the K-NN as techniques to measure the quality of learning.

In our search studies, we distinguish two sets of experiments dedicated mainly to the performance evaluation of the proposed method. The first set is manifested by the manual division data into two subsets; one set for learning (80% of the corpus) and a second a distinct set for the test (20% of the corpus). This set allows presenting the evaluation results of the learning and testing phases. The second experimentation set is automatically carried out, using cross validation that allows presenting the results of the ranking phase (test).

The following section consists in presenting the results obtained from the evaluation of our system. It is composed of two parts: the first part presents the results of the evaluation of learning and the second presents the results of the evaluation of the ranking result documents.

Experimental Set 1: manual division. In this section, we present two types of the obtained results: those obtained after learning and those resulting from the projection

of the test corpus on the prediction model. Thus, the used evaluation measures are accuracy, recall, precision, F-measure and kappa.

Learning results. In the context of the evaluation by manual division of the corpus and using decision tree algorithms, SVM and KNN, we obtained the results presented in table 2. By referring to this table, it therefore appears obvious that the results of our learning method are very interesting. Indeed, in the case of the KNN algorithm, the recall is in the order of 74.6% whereas precision is equal to 78.1%, hence, the F-measure is equal to 72.1%. Likewise, we obtained an accuracy of 74.6%. Finally, we have achieved a kappa degree of agreement between prediction and supervision which is equal to 0.56.

Finally, in the case of the algorithm of the decision tree, the recall is of the order of 77% while precision is equal to 77.6%, hence, the F-measure is equal to 76.5%. Likewise, we obtained an accuracy of 77%. Finally, the achieved degree of agreement between prediction and supervision (kappa) is equal to 0.61.

Table 2. Experiment No. 1: Evaluation results of the learning phase by manual division based on the SVM, KNN and the decision tree.

	Accuracy	**Recall**	**Precision**	**F-measure**	**Kappa**
SVM	47.4%	74.4%	54.8 %	42.3 %	0.07
KNN	74.6 %	74.6%	78.1%	72.1 %	0.56
Decision tree	77 %	77%	77%	76.5%	0.61

Ranking result. This phase is to use the prediction model obtained from the learning phase to classify new documents. In the context of the evaluation using manual division of the corpus as well as the following algorithms; the decision tree, the SVM and the KNN, we obtained the results presented in table 3. According to this table, it appears that the results of our ranking method are interesting. Indeed, in the case of the algorithm of the decision tree, the recall is in the order of 66.1 % while precision is equal to 72 %, therefore, the F-measure is equal to 67.3 %. Similarly, the obtained accuracy is 66 %. Finally, the degree of agreement archived between prediction and supervision (kappa) is equal to 0.41.

Table 3. Experiment 1: evaluation results of the ranking phase by manual division of the corpus based on the SVM, KNN and the decision tree.

	Accuracy	**Recall**	**Precision**	**F-measure**	**Kappa**
SVM	51.8%	51.9 %	60.6 %	48.5 %	0.16
KNN	68.7 %	60.2%	54.8%	55.9%	0.17
Decision tree	66 %	66.1%	72%	67.3%	0.41

Experimental Set 2: cross-validation. To classify new documents, the proposed ranking method consists in using the classification model obtained during the learning

phase. Therefore, the evaluation of the ranking method is to evaluate the predictive model with new documents. On the other hand, the evaluation measures that we have used are the same as those of the evaluation of the learning model, namely, accuracy, confusion matrix, recall, precision, F-measure and kappa.

In the evaluation context using cross-validation (K-fold) with K = 26, the decision tree, the SVM and the KNN algorithms, we obtained the results presented in table 4. From this table, it appears that the results of our ranking method are interesting. Indeed, in the case of the SVM algorithm, the recall is in the order of 60.6 % whereas precision is equal to 45.6 %, hence, the F-measure is equal to 46.1%, besides, an accuracy of 60.5% is obtained. Finally, we can say that the achieved kappa degree of agreement between prediction and supervision is equal to 0.11. Finally, in the case of the algorithm of the decision tree, the recall is in the order of 61.4 % while precision is equal to 58 %, consequently, the F-measure is equal to 59.2 %. Likewise, we obtained an accuracy of 61.4 %. Finally, it can be noted that we have achieved a degree of agreement between prediction and supervision (kappa), which is equal to 0.24.

Table 4. Experiment No. 2: Evaluation results of the ranking phase using the cross-validation method k-fold based on the SVM, KNN algorithms and the decision tree.

	Accuracy	**Recall**	**Precision**	**F-measure**	**Kappa**
SVM	60.5%	60.6 %	45.6 %	46.1 %	0.11
KNN	60.1 %	60.2%	54.8%	55.9%	0.17
Decision tree	61.4 %	61.4%	58%	69.2%	0.24

The discussion of the learning results, using cross validation shows that the decision tree increases the performance of our learning model. For this reason, in the context of our ranking method, we adopted the algorithm of the decision tree to build the predictive model which is also used to classify new returned documents for a query submitted by the same user.

Similarly, we performed a set of learning experiments with the user profile (which means that we have integrated the learning criteria linked to the user profile in the learning model) and a series of experiments without the user profile (that is to say, we eliminated the user profile-related learning requirements from our learning model).

Table 5. Evaluation of the learning outcomes by integrating the user profile and learning outcomes without the use of the user profile.

	Accuracy	**Recall**	**Precision**	**F-measure**	**Kappa**
Learning with profile	61.4 %	61.4 %	58 %	69.2 %	0.24
Learning without profile	40.4 %	41 %	35 %	37.7 %	0.11

As shown in table 5, we found, in all cases, that learning by means of the profile has given better results than without it. Indeed, the accuracy of learning by means of the profile is equal to 61.4%, while that without it is about 40.4%. This proves the contribution of the hybrid user profile in our ranking system.

4.4 Comparison to Baseline Methods

On the other hand, we have also experimentally compared our "SPIRAL" contribution to the method of the search engine "Lucene" (a baseline method in our case). In fact, Lucene uses a model which is derived from Boolean model. Thus, Lucene method is a method without profile that is to say without personalization of IR.

Table 6. Performance gain of personalized search (precision and MAP Measures).

Precision	baseline method	SPIRAL	MAP	baseline method	SPIRAL
%P10	9	20	**%MAP5**	7	14
%P20	10.1	16.9	**%MAP10**	5	13
%P30	3.2	10	**%MAP15**	6	15
%P50	6	9.9	**%MAP**	6	14

Calculation of Precision Average. The results for the SPIRAL system (with hybrid profile) are better than those of the baseline method (Table 6). Indeed, the precisions P10, P20, P30 and P50 of the SPIRAL are better than the one in the baseline method. As a conclusion, we have demonstrated that personalizing the IR showed better results with a hybrid profile than IR with a base line method.

Calculation of MAP (Mean Average Precision). We notice that the results obtained with the SPIRAL system are better than those obtained with the baseline method (Table 6.). Moreover, the MAP5, MAP10 and MAP15 for SPIRAL are better than those of the baseline method. Indeed, SPIRAL system shows all these performances for the first 15 documents by MAP15 = 15 and its MAP is better than the Lucene system by MAP15 = 6. Similarly, we can see that the IR showed better results with hybrid profile (personalization) than with a baseline method.

4.5 Discussion of Results

In a first set of experiments, we have divided our corpus (20,000 documents) in two corpora, namely a training corpus (16,000 documents) and a test corpus (4,000 documents). The results obtained by exploiting the algorithm of the decision tree, when evaluating the learning phase, are very interesting with an accuracy equal to 77%. Thus, the predictive model obtained from the learning phase is a performing model that has interesting results from our ranking system with accuracy equal to 66%.

In a second set of experiments, we used the algorithm of the decision tree when evaluating the ranking phase. It was found that the obtained results are interesting. Indeed, we have obtained an accuracy of 61.4%. It is emphasized that we had 61.4% as recall and 58 % as precision; hence the F-measure is equal to 69.2%. Similarly, it can be said that we have achieved a degree of agreement between prediction and supervision (kappa) equal to 0.24.

On the one hand, the obtained recall rate is explained by the ability of our learning model to return a large number of relevant documents among all the relevant ones in the corpus. This is explained by the contribution of the hybrid user profile in the process of finding relevant documents.

Furthermore, through our hybrid profile, the ranking system helped to return a large number of relevant documents among all the ones proposed by the system, which explains the precision rate of 58 %.

Nevertheless, the Kappa value of 0.24 indicates that the proposed ranking system allows a relatively medium degree between prediction (predicted class) and supervision (real class).

Also, it is observed that the length of the query has a relatively direct impact on the results of our system. Indeed, it was found that if the number exceeds four terms without expansion and the expansion process adds to each term at least three other concepts from the hybrid user profile, then, we'll get at least 12 terms in the enriched query. This will generate a lot of noise in the document search process and therefore, more irrelevant documents.

Similarly, after a comparison of our "SPIRAL" contribution against a baseline method, we can see that the personalization of IR by %MAP = 14 showed better results with the hybrid profile than IR with a baseline method by %MAP = 6.

Concerning the learning criteria, we emphasize that we first adopted classical criteria (the first category and the fourth category of the criteria) used by the majority of the studies on the IR. Secondly, we decided to add user profile criteria (the third category of criteria) and semantic criteria (the second category of criteria). This enabled us to further improve the results that passed for the P10 precision rate from 9% to 20% and the percentage of the MAP average from 6% to 14%.

As a conclusion, one of the strengths of the proposed method of RIP has five aspects:

- The proposed method is interesting because it is more user-oriented progressively adapts to the evolution of his profile and his knowledge.
- Learning is performed for each user apart from what proves the personalization aspect characterizing the method.
- The contribution of the hybridization of the user profile (the conceptual and multidimensional representation) to the mechanisms of the query expansion and the ordering of the documents restored by a search engine.
- The positive impact of the semantic learning criteria (based on information from semantic resources) and the criteria related to the user profile (represented by hierarchies of concepts) on the performance of our RIP system.
- Integrating the user profile in all the levels of the PIR process.

5 Conclusion and Prospects

In this work, we focused on the method of ranking documents that we proposed as part of a personalized information retrieval system. The proposed personalized learning to rank method is based on the integration of the user profile into the learning criteria and the proposed ranking function. The representation of the user profile (hybrid approach) in our method is based on the extraction of semantic relationships found in ontologies (AWN and Amine AWN) i.e. synonymy, hyperonymy and hyponymy.

To ensure the achievement of the ranking method, we used a learning model that exploits the user' explicit relevance judgments. This consists in asking the user to assign a relevance class to a document which reflects the importance of the document with respect to the user's needs. In a second phase, we projected these judgments on criteria related to a document, a query and a profile. This projection helps build a predictive model that can discern relevant documents meeting the profile at the user's query. The predicted model will then be used in the ranking phase to classify other document results from a new query submitted by the same user.

Similarly, we have devoted a part of this article to describe the implementation of a document ranking system of Arabic entitled "SPIRAL". To evaluate the proposed method, we have used a corpus of 30,550 Arabic texts that covers topics related to the field of «علوم طبيعية» "natural sciences". The results of our evaluation ranking system prove the performance of the latter. In fact, we noticed that the results of our ranking method with the cross-validation model (K-fold with k = 26) are interesting. Indeed, the F-measure is in the order of 59.2%. Similarly, we obtained 61.4% as an accuracy rate. Finally, it can be noted that we have achieved a degree of kappa agreement between prediction and supervision equal to 0.24.

Thus, the accuracy of learning by means of the profile is equal to 61.4%, while that without it is about 40.4%. In addition, we note that the semantic learning criteria related to the user have a positive impact on the performance of SPIRAL system. This justifies our choice of the integration of the hybrid user profile into the learning criteria.

At this stage, we can distinguish several research perspectives. Therefore, in the short term, we can choose evaluating the user profile by studying the impact of the number of relevant documents in building the profile, the ranking parameter results and the depth of the hierarchy of the concept profile in improving the search results. Similarly, we intend to build a profile based on search history and compare it with our hybrid profile.

It is emphasized that the evaluation method of learning to rank was made using our own corpus "WCAT" and according to a simulation scenario of TREC research sessions. In order to validate the effectiveness of our method in a real research environment, our outlook in the medium and long term, is to evaluate this method using data from a log file of a search engine.

References

1. Jenifer, K., Yogesh Prabhu, M., Gunasekaran, N.: A survey on web personalization web approaches for efficient information retrieval on user interests. Journal of Recent Research in Engineering and Technology, 2, 34–39 (2015)
2. Jayanthi, J., Rathi, DR. S.: Personalized web search methods – A complete review. Journal of Theoretical and Applied Information Technology (JATIT). 62, 685–697 (2014)
3. Akhila, G.S., Prasanth, R.S.: A Survey on Personalized Web Search. Journal of Advanced Research Trends in Engineering and Technology (IJARTET). Vol.2 (2015)
4. Safi, H., Jaoua, M., Belguith, L.: PIRAT: a Personalized Information Retrieval system in Arabic Texts based on a hybrid representation of a user profile. In: 21st International Conference on Applications of Natural Language to Information Systems (NLDB), pp. 326–334, LN CS, Springer, Royaume-Uni (2016)
5. Shou, L., Bai, H., Chen, K., Chen, G.: Supporting Privacy Protection in Personalized Web Search. IEEE Trans. on Knowledge and Data Engineering. 26, 453–467 (2014)
6. Cossock, D., Zhang, T.: Subset ranking using regression. In: 19th annual conference on learning theory, pp. 605–619. Springer-Verlag, Heidelberg, Berlin (2006)
7. Chapelle, O., Keerthi, S. S.: Efficient algorithms for ranking with svms. Information Retrieval, 13, 201–215 (2010)
8. Xia, F., al.: Listwise approach to learning to rank: theory and algorithm. In: 25th International conference on machine learning, pp. 1192–1199. ACM, New York, NY, USA (2008)
9. Taylor, M.: Softrank: optimizing non-smooth rank metrics. In: International conference on web search and web data mining, pp. 77–86 (2008)
10. Zoeter, O., Taylor, M., Snelson, E., Guiver, J., Craswell, N., Szummer, M.: A decision theoretic framework for ranking using implicit feedback. In: SIGIR 2008 workshop on learning to rank for information retrieval (2008)
11. Xu, J., Li, H.: Adarank: a boosting algorithm for information retrieval, In: 30th annual international ACM SIGIR conference on research and development in information retrieval, pp. 391–398, ACM, New York, NY, USA (2007)
12. Yue, Y., Finley, T. Radlinski, F., Joachims, T.: A support vector method for optimizing average precision. In: 30th annual International ACM SGIR conference on research and development in inform. retrieval, pp. 271–278, ACM, New York, NY, USA (2007)
13. Cao, Z., Qin, T., Liu, T.-Y., Tsai, M.-F. Li, H.: Learning to rank: from pairwise approach to listwise approach. In: 24th International conference on machine learning, pp. 129–135, Technical Report, Corvallis, OR, ACM, MSR-TR-2007-40 (2007)
14. Chapelle, O., Zhang, Y.: A dynamic bayesian network click model for web search ranking, In: 18th International conference on World Wide Web, pp. 1–10, New York, USA, ACM (2009)
15. Liu, T.-Y.: Learning to rank for information retrieval. LNCS, Berlin Heidelberg (2011)
16. Baccini, A., Déjean, S., Lafage, L., Mothe, J.: How many performance measures to evaluate information retrieval systems? Knowledge and Information Systems, 30, 693–713 (2012)
17. Baziz, M. : Indexation conceptuelle guide par ontologie pour la recherché d'information. Ph.D. dissertation, Université de Paul Sabatier, Paris (2005)
18. Mezaour, A. D. : Recherche ciblée de documents sur le web. Ph.D. dissertation, Université Paris-Sud, Paris (2005)
19. Abouenour L., et al.: Improving Q/A using Arabic Wordnet. In: International Arab Conference on Inform. Technology (ACIT), Zarqa Private University, Jordan (2008)
20. Xu, J., Fraser, A., Weischedel, R.: Empirical Studies in Strategies for Arabic Retrieval. In: Ann. ACM Conference on Research and Development in Inform. Retrieval: In: 25th annual

International ACM SIGIR conference on Research and development in inform. Retrieval, 11, 269–274 (2002)
21. Bessou, S., Saadi, A., Touahria, M.: Vers une recherche d'information plus intelligente application à la langue arabe. In: 1st International Conference on Inform. System and Economic Intelligence (SIIE), Hammamet, Tunisia (2008)
22. Hammo, B.: Effectiveness of Query Expansion in Searching the Holy Quran. In: International Conference on Arabic Language Process. (CITALA), pp. 18–19, Morroco (2007)
23. Zaidi, S., Laskri, M.: Expansion de la requête Arabe sur le réseau internet. In: Barmajiat CSLA: Les applications logicielles en arabe: Pas vers l'e-gouvernement, Alger (2007)
24. Farag, A., Andreas, N.: AraSearch: Improving Arabic text retrieval via detection of word form variations. In: International Conference on Information System and Economic Intelligence (SIIE), Hammamet, Tunisia (2008)
25. Abderrahim, M.-A. Abderrahim, M.-E., Chikh, M.-A.: Using Arabic WorldNet for Query Expansion in Information Retrieval System. In: IEEE 3th International Conference on Web and Information Technologies, Marrakech, Morocco (2010)
26. Abderrahim, M-A. : Utilisation des ressources externes pour la reformulation des requêtes dans un système de recherche d'information. The Prague Bulletin of Mathematical Linguistics, PBML, 99, 87–99 (2013)
27. Kleinberg, J.: Authoritative sources in a hyperlinked environment. Journal of the ACM, 5, 604–632 (1999).
28. Kohavi, R., Quinlan, R.: Decision tree discovery. Handbook of Data Mining and Knowledge Discovery, University Press, pp. 267–276 (1999)
29. McLachlan, Do., K.-A., Ambroise., C.: Analyzing Microarray Gene Expression Data. NJ, USA (2004)
30. Arabic Wordnet, http://wwww.globalwordnet.org/AWN/ AWNBrowser.html
31. Amine Arabic Wordnet, http : //amine-platform.sourceforge.net

ISSN 1870-4069

Arabic Cooperative Answer Generation via Wikipedia Article Infoboxes

Omar Trigui[1], Lamia Hadrich Belguith[2], Paolo Rosso[3]

[1] ANLP Research Group, MIRACL Lab, ISAEG, University of Gafsa, Tunisia

[2] ANLP Research Group, MIRACL Lab, FSEG, University of Sfax, Tunisia

[3] PRHLT research center, Universitat Politècnica de València, Spain

omar.trigui@fsegs.rnu.tn, l.belguith@fsegs.rnu.tn, prosso@dsic.upv.es

Abstract. The typical question-answering system is facing many challenges related to the processing of questions and information resources in the extraction and generation of adequate answers. These challenges increase when the requested answer is cooperative and its language is Arabic. In this paper, we propose an original approach to generate cooperative answers for user-definitional questions designed to be integrated in a question-answering system. This approach is mainly based on the exploitation of the semi-structured Web knowledge which consists in using features derived from Wikipedia article infoboxes to generate cooperative answers. It is globally independent of a particular language, which gives it the ability to be integrated in any definitional question-answering system. We have chosen to integrate and experiment it in a definitional question-answering system dealing with the Arabic language entitled DefArabicQA. The results showed that this system has a significant impact on the approach efficiency regarding the improvement of the quality of the answer.

Keywords: Natural language processing, natural language generation, data extraction and integration, web knowledge, question answering system, cooperative answers, Arabic language.

1 Introduction

In this paper, we propose an approach that allows the generation of definitional cooperative answers using semi-structured Web knowledge bases. This approach is designed to be integrated in a definitional question-answering system. Before presenting the approach in detail, we will specify the general context, the motivations, the challenges to face, and the objectives to be achieved through this approach.

A question-answering system presents the intersection of several domains, such as information retrieval, information extraction, and natural language processing. Its typical interest is to make access to information toward information resources of large sizes and with heterogeneous, fast and smooth structures. A huge progress in results is achieved. This has been proved by organizing a series of competitive workshops dealing with the question-answering track by international conferences, such as TREC[1], CLEF[2] and NTCIR[3]. Different types of questions were dealt with these competitive workshops such as the complex one like "definition" question type. Typically, a question of the type "definition" is the one that asks about important information related to a fact, a person, an organization or an event. The adopted answer form to a definition question by competitive workshops is a list of information nuggets (i.e. a set of pieces of important informational texts) [Voorhees, 2003].

However, this answer form remains modest at the level of the structure and harmonization of information compared to what is expected by a user as an expected defining answer. This one presents for us a motivation to enhance this form of answer and to propose a new answer form reflecting high level coherence information and well-structured definition answer.

Fig. 1. An extract of a Wikipedia article entitled Mark Zuckrberg[4].

Many challenges arise when we get into dealing with the question-answering systems. We cite two aspects that are very important for their functioning. The first can handle information resources that are not well-formed, while the second one can find out the type of information looked up behind the question and deduce the details of the expected answer. These challenges increase when we deal with the "definition"

[1]http://trec.nist.gov/
[2]http://clef2015.clef-initiative.eu/
[3]http://research.nii.ac.jp/ntcir/index-en.html
[4]http://en.Wikipedia.org/wiki/Mark_Zuckerberg

question type and with a developed answer form for the expected defining answer adding the specificities of a language characterized by their low resources.

In this paper, we propose an approach which permits to generate a cooperative answer in the form of a paragraph designed to be integrated in a definition question-answering system dealing with low resource languages. The concept of this approach is based on the exploitation of semi-structured Web knowledge bases and specifically Wikipedia article infoboxes. We have chosen to exploit Wikipedia as it is a much large semi-structured Web knowledge base which contains more than 30 million articles in 287 languages. On the one hand, its website is the fifth most visited website in the world with 18 billion visitors[5], which can prove the confidence of their information and their wide coverage of topics, on the other hand. Among the components of a Wikipedia article, there is a one entitled infobox which contains a summary of important information relative to the main subject dealt with in a given Wikipedia article. These pieces of information are often located in a formatted box at the top of a Wikipedia article.

Figure 1 shows an example of an extract of a Wikipedia article entitled Mark Zuckrberg, while figure 2 shows an example of an infobox extracted from the Wikipedia article entitled "Jimmy Wales".

Fig. 2. An example of a Wikipedia article infobox entitled Jimmy Wales[6].

[5]http://en.Wikipedia.org/wiki/Wikipedia
[6]http://en.Wikipedia.org/wiki/Jimmy_Wales

2 State of the Art

In this section, we will present the main research studies based on the exploitation of Wikipedia as a Web knowledge base, as well as the major studies on generating answers in question-answering tasks. The exploitation of Wikipedia as a Web knowledge base has been introduced in various research studies addressing information retrieval, information extraction, construction of multilingual corpus and automatic translation [Lopez et al., 2011]. Among these research studies, we can mention those of Bizer et al. [2009] and Yahia & Salhi [2014], which are based on the exploitation of Wikipedia as a Web knowledge base in the information retrieval field. Bizer et al. [2009] made a great effort to extract structured information from Wikipedia and make it accessible to the Web. The resulting DBpedia knowledge base currently describes more than 2.6 million entities. However, Yahia & Salhi [2014] used Wikipedia as a knowledge base for the categorization of documents. Other research studies exploiting Wikipedia article infoboxes and addressing information retrieval appeared in a set of workshops. One of these interesting workshops is KBP[7] "Knowledge-Base Population" which has been organized by the TAC[8] conference since 2009 [Ji and Grishman, 2011; Surdeanu, 2013].

Regarding question-answering systems, further research studies based on Wikipedia have been designed. We can cite those which exploit Wikipedia as a knowledge base like [Trigui et al. 2010a; Brzeski&Boi´nski, 2014; Yang et al. 2014] and [Ryu et al. 2014]. The research of Trigui et al. [2010a] exploited Wikipedia article contents through a Web search engine in an Arabic definition question answering system to build a specific information resource relative to each given question. On the other hand, Yang et al. [2014] proposed a method to build a robust knowledge resource based only on semantic associations automatically extracted from Wikipedia. The obtained knowledge resource was designed to be integrated in a question-answering system. On their part, Breski & Boinski [2014] proposed a method which is based on associations between Wikipedia articles to answer factual questions. Moreover, Ryu et al. [2014] proposed a method to categorize the Wikipedia structures into article contents, infoboxes, category structures, article structures and redirection links. These Wikipedia structure categories were designed to be used as a rich knowledge resource for factual question-answering systems. Other research studies dealing with question-answering systems exploited Wikipedia for the validation of answers [Buscaldi & Rosso, 2006; Cui et al. 2007].

We have cited examples of research studies based on Wikipedia exploitation as a Web knowledge base in the domains information retrieval and question-answering. Here, we cover the details of the major studies on generating answers in question-answering. Many research studies have dealt with the question-answering systems, but only a few of them have addressed the answer-generation step beyond the answer extraction step [Voorhees, 2004; Dang et al., 2007]. Typically, the answer-generating step permits to generate an answer where there is more than one possible answer or no

[7]http://pmcnamee.net/kbp.html
[8]http://www.nist.gov/tac/tracks/index.html

answer found in the data resources [Benamara & Saint-Dizier, 2004]. It would be an indirect answer to the user's question and more helpful than the direct one [Corella & Lewison, 2009]. This form of answer is entitled 'a cooperative answer' [Benamara, 2004]. Most of the research studies dealing with cooperative answer generation are based on integrating knowledge representation and reasoning mechanisms. We can mention, for example [Prager et al. 2003; Benamara & Saint-Dizier, 2004] and [Lupkowski & Leszczyńska-Jasion, 2014]. Prager et al. [2003] proposed to answer a definition question by gathering answers to factual questions derived from the given definitional question. As for Benamara & Saint-Dizier [2004], they proposed an approach that enables to answer factual questions in French by cooperative answers dealing with the tourism domain. On their part, Lupkowski & Leszczyńska-Jasion [2014] described a system designed to generate cooperative answers based on inferential erotetic logic concepts.

We will focus now on the study of research studies addressing the Arabic Web knowledge base in information retrieval and question-answering fields. In the literature, research studies dealing with the Web knowledge base in information retrieval and addressing the Arabic language have witnessed a growing interest during the last few years [Ezzeldin&Shaheen, 2012; AlZoghbyaa et al., 2013]. We can mention in this respect the research studies of Beseiso et al. [2011], Al-Zoghby and Shaalan [2015], and Al-Bukhitan et al. [2014]. These research studies have shared the goal of facilitating the search and the access to information adopting the Semantic Web technology in information retrieval.

Beseiso et al. [2011] proposed a new framework intended to add a semantic Web layer to the current Web-based applications in order to improve the searching and linking processes. However, Al-Bukhitan et al. [2014] proposed an automatic annotation tool that supports the semantic annotation of Arabic Web documents for semantic search engines. A promising performance was achieved by this automatic annotation tool. On the other hand, Al-Zoghby and Shaalan [2015] proposed a semantic search approach applied to Arabic Web content which is based on the Vector Space Model. It consists in locating Web contents that are semantically related to the query's concepts rather than relying on the exact matching with keywords in queries.

In spite of the efforts made in the Arabic language to adopt Web knowledge in information retrieval, there is a lack of research studies adopting and exploiting Web knowledge in question-answering. This can be explained by the structure of the Arabic Web knowledge bases which make it particularly difficult to handle the automatic processing challenges of the Arabic language properties by question-answering systems. We mention two research studies which used the Web as a knowledge base [Trigui et al, 2010a; Hasanain et al. 2014]. The first is based on Web search engines to construct specific resource knowledge of snippets for each given question in a definition question-answering system. The second is based on Twitter to construct corpus constituted of millions of tweets for its system [Hasanain et al., 2014].

Apart from that, there are various research studies dealing with Arabic question-answering systems based on closed corpus of documents [Ezzeldin & Shaheen, 2012; Shaheen & Ezzeldin, 2014]. With respect to generating answers in the Arabic language, we have to mention that, to our knowledge, there are practically no research

studies that addressed this issue. Most prominent Arabic question-answering systems return the answers in a paragraph form, such as Hammo et al. [2004] or in a list of information nuggets, as in the case of Trigui. [2011]; Badawy et al. [2011]; Fareed et al. [2014] and Kurdi et al. [2014].

To sum up, we have cited research studies dealing with Wikipedia as a Web knowledge base in various manners for question-answering systems. Among these research studies, we have mentioned question-answering systems which generate cooperative answers based on integrating knowledge representation and reasoning mechanisms. In addition, we have shown the important research studies addressing Arabic question-answering systems and their various forms of answers. As for our approach, it is meant to exploit the wealth of Wikipedia information, and especially Wikipedia article infoboxes, for generating answers to definition questions in the context of a definition question answering system.

3 Approach for Generating Cooperative Answers

In this section, we detail the proposed approach to generate cooperative answers to definitional questions for question-answering systems. It is based on a part of the semi-structured Web knowledge base presented by the Wikipedia article infoboxes. It can be integrated in each question-answering system dealing with definitional questions, independently of its particular language. This approach involves three main tasks: the infobox class generation, the cooperative answer pattern generation and the cooperative definition answer generation (see figure 3). Hereafter, we will detail these three tasks.

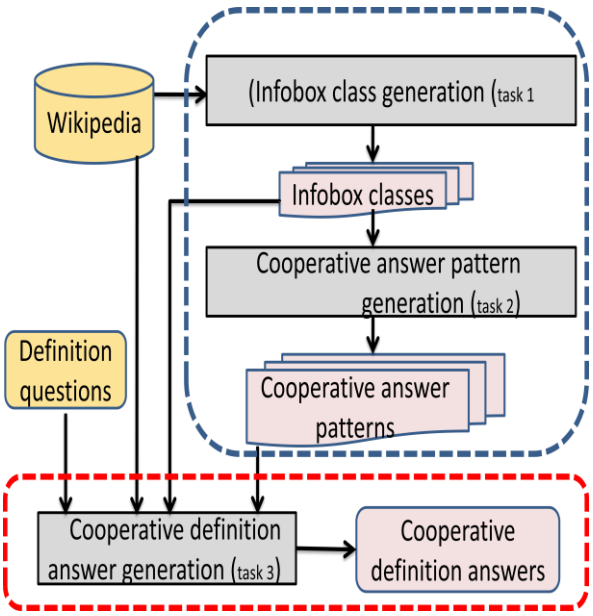


Fig. 3. Main tasks of the proposed approach.

3.1 Infobox Class Generation (Task 1)

The infobox class generation task includes a set of three sub-tasks, namely: Wikipedia article selection, infobox exploitation and infobox grouping, as mentioned in figure 4. These sub-tasks are based on the information in the Wikipedia article infoboxes. It exploits the hypothesis that a given Wikipedia article infobox shows a resource of specific and relevant information relative to a definite named entity in a given language.

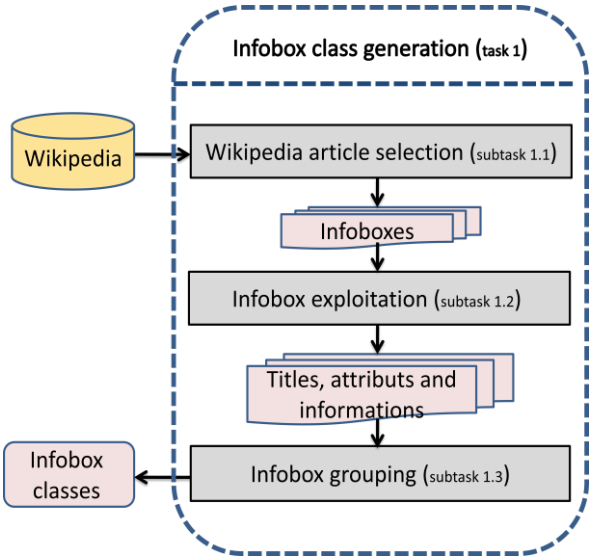


Fig. 4. The subtasks of the infobox class generation task.

Wikipedia Article Selection (sub-task 1.1). Wikipedia contains hundreds of thousands of articles in different languages where each article represents a separate Web page describing a definite named entity (e.g. an event, a person, an organization or a concept). This sub-task consists of browsing Wikipedia for a given language in order to select Wikipedia articles containing infoboxes. The number of Wikipedia articles varies from one language to another. It reaches more than four million items in the English language and around three hundred thousand items in the Arabic language, as an example[9].

Infobox Exploitation (sub-task 1.2). This sub-task consists in extracting infobox features (i.e. attributes, title and information) from Wikipedia articles containing Infoboxes. Figures 2 and 5 present two infoboxes containing summary of specific information respectively describing Jimmy Walles and Bill Gates (two Wikipedia article titles). Table 1 shows attributes extracted from two infoboxes, respectively, enti-

[9]http://en.wikipedia.org/wiki/Wikipedia#Language_editions

tled “Jimmy Wales” and “Bill Gates”. The first infobox attributes are “born”, “residence”, “other names”, “alma mater”, “occupation”, “title”, “successor”, “board member of”, “spouse”, “awards” and “Website” while the attributes of the second one are “born”, “residence”, “alma mater”, “occupation”, “active years”, “net worth”, “board member of”, “spouse”, “children”, “parents” and “Website”. These examples confirm that each infobox presents a resource of the basic information describing a precise named entity which can be a person, an organization, a date, a location, or an event. In our case, infobox attributes present the most important features of an infobox.

Fig. 5. Example of an infobox entitled Bill Gates[10].

Table 1. Attribute list extracted from the infoboxes entitled “Jimmy Wales” and “Bill Gates”.

Jimmy Wales infobox attributes		**Bill Gates infobox attributes**	
Born	***boardmember of***	***Born***	***boardmember of***
Residence	*successor*	***Residence***	***Spouse***
other names	***spouse***	***alma mater***	*Children*
alma mater	*awards*	***occupation***	*Parents*
Occupation		*active years*	*Signature*
Title	***website***	*net worth*	***Website***

[10]http://en.wikipedia.org/wiki/Bill_Gates

Grouping Infoboxes (subtask 1.3). The sub-task consists in grouping infoboxes in a class of infoboxes according to a minimum rate of similarity between their attributes (those in common). For each infobox class, we consider the following features: a label (i.e. a semantic category), which is typically a hyperonymy of the included infoboxes in this class, as well as a set of attributes shared by the various semantically closest infoboxes.

Table 2. The infobox class features entitled "entrepreneur".

Semantic category	
Entrepreneur	رجل الأعمال
Attributes	
Born	ولد/ت
Residence	إقامة
Alma mater	جامعة
Occupation	ت/يشغل منصب
Board member of	عضو مجلس إدارة
Spouse	شريك حياته/ها
Website	موقع الويب

Table 3. The infobox class features entitled "political party".

Semantic category	
Political party	حزب سياسي
Attributes	
Founded by	تأسس من طرف
Foundation year	سنة التأسيس
party leaders	قادة الحزب
Ideology	الإيديولوجية المتبعة
Location	الموقع
the leader	القائد
number of deputies	عدد النواب
leadership center	مركز القيادة
Website	موقع الويب

The result is of this task is a set of infobox classes where each infobox class contains the semantically closest infoboxes from the set of Wikipedia article infoboxes. A semantic category is attributed to each infobox class by a person who would be a reader of the international newspapers and native speaker of the respective language of the infoboxes.

Typically, these semantic categories provide titles for infobox classes. Tables 2, 3 and 4 show three infobox classes with their respective features mentioned in English and in Arabic. For example, the infobox class cited in table 2, which has the semantic category 'entrepreneur', is the result of the grouping of the semantically closest infoboxes, such as the infoboxes entitled Jimmy Wales and Bill Gates relatively in figures 2 and 5. Their shared attributes, which are mentioned in bold in table 1, constitute the attributes of this infobox class, such as "born", "residence", "alma mater", "occupation", "board member of", "spouse" and "Website".

Table 4. The infobox class features entitled "sports team".

Semantic category	
Sports team	فريق رياضي
Attributes	
full name	الاسم الكامل
Nickname	الكنية
Founded	أسّس
Stadium	الملعب
League	الدوري
Coach	المدرب
Website	موقع الويب

3.2 Cooperative Answer Pattern Generation (Task 2)

The cooperative answer pattern generation task consists in building a set of cooperative answer patterns where each one is taken as a skeleton of a definition answer. It consists in exploiting infobox class features. The sub-tasks are involved in this task are, respectively, composing skeletons and formulation of cooperative patterns (see figure 6).

Composing skeletons (sub-task 2.1). This subtask consists in generating a pattern answer skeleton for each infobox class using all its attributes. The attributes of each infobox class are taken in order, one by one, in the composition of the respective pattern answer skeleton in the first step. In the second step, a blank reserved for respective information is added after each attribute in the composing pattern answer skeleton.

Formulation of cooperative answer patterns (sub-task 2.2). This sub-task consists in checking each composed pattern answer skeleton and adding specific punctuation marks (i.e. a comma, a full stop, etc.) after each blank reserved for information in the patterns. The choice of the respective punctuation marks is made by a linguistic expert.

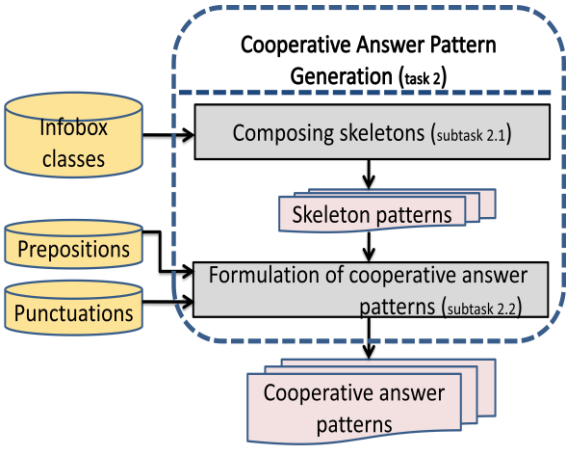


Fig. 6. Subtasks constituting the cooperative answer pattern generation task.

Table 5. The cooperative answer pattern relative to the semantic category "entrepreneur".

A COOPERATIVE ANSWER PATTERN
<VALUEQUESTION FOCUS> + IS +{BORN}+ IN + <VALUE> + . + HIS {RESIDENCE}+ IS AT + <VALUE> + . + HIS {ALMA MATER }+ IS AT THE + <VALUE> +.+ HIS + {OCCUPATION}+ IS/ARE + <VALUE> +,+ {BOARD MEMBER OF} + <VALUE> +. +HIS + {SPOUSE} + IS + <VALUE>+. +HIS + {WEBSITE} + IS + <VALUE>+.
<المعلومةموضوع سؤال > {ولد/ت} + في + > المعلومة> +. + {تقيم/يقيم} + في+<المعلومة> +.+ {متخرج/ة من } + <المعلومة> +.+ {يشغل منصب} + <المعلومة> +،+ { عضو مجلس إدارة } + <المعلومة> +. + {شريك حياته/ها} + هو/هي + <المعلومة> +.+{ موقع الويب الخاص+}هو+<المعلومة>.+

Table 6. A cooperative answer pattern relative to the semantic category "political party".

A COOPERATIVE ANSWER PATTERN
<VALUE QUESTION FOCUS> + IS + {FOUNDED BY} + <VALUE>+ . + THE + {FOUNDATION YEAR} + IS + <VALUE>+ . + THE + {PARTY LEADERS}+ ARE + <VALUE> + . + IT + {IDEOLOGY} + <VALUE> +.+ THE {LOCATION} + IS + <VALUE> +. THE + {LEADER}+ IS + <VALUE> +. THE + {NUMBER OF DEPUTIES} + IS + <VALUE> + AND THE + {LEADERSHIP CENTER}+ IS + <VALUES> +.
<المعلومةموضوع سؤال > {تأسس من طرف } + <المعلومة> +. +{سنة التأسيس} + كانت في+ <المعلومة> +. { قادة الحزب } + هم+<المعلومة> +. + {الإيديولوجية المتبعة} + هي +<المعلومة> +.+ { الموقع} + هو +<المعلومة> +.+ {القائد } + هو/هي +<المعلومة> +.+ { عدد النواب } + هو+ <المعلومة> +.+ { مركز القيادة } + هو+ <المعلومة> +.

These two sub-tasks enable to generate a cooperative answer pattern for each infobox class. Tables 5 and 6 show two examples of cooperative answer patterns in English and in Arabic. The first is related to the infobox class having the semantic

category "entrepreneur", while the second is related to the infobox class having the semantic category “political party”. The texts inside the curly braces are these attributes while the annotation "<value>" means the respective information of the associated attribute. The annotation “<value $_{\text{Focus question}}$>” means the focus of the given question.

3.3 Generating Answers (Task 3)

The answer generation task is composed of three sub-tasks. It consists in generating cooperative definition answers to definition questions. The first sub-task consists in extracting the main named entity of the given definitional question (i.e. the question focus). The second one consists in selecting the adequate cooperative answer pattern relative to the given question, while the last sub-task consists in filling the blank of the selected answer pattern with the respective information. We will detail these sub-tasks one by one below.

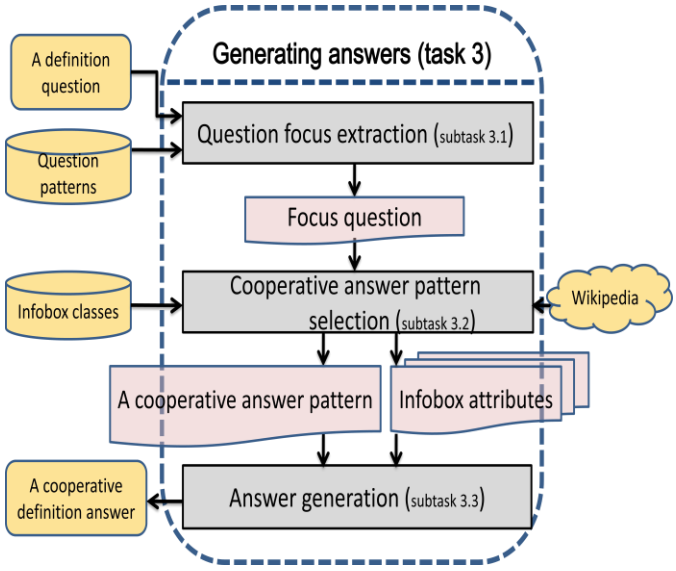


Fig. 7. Sub-tasks constituting the generating answer task.

Question Focus Extraction (subtask 3.1). It consists in identifying the main named entity presenting the interest subject of a definition question (i.e. the question focus). The identification is based on a set of lexical patterns designed for definition questions. Table 7 shows examples of typical lexical patterns to identify question focuses from definition questions in English and in Arabic.

Table 8 presents an example of a definition question as well as its respective lexical pattern of definition questions in English and in Arabic. The definition question is «Who is Steve Chen? » and its respective question focus is « Steve Chen » identified by the following lexical pattern: “Who + be + <a question focus>+ ?”.

Table 7. Lexical patterns of the definition questions in English and in Arabic [Benajiba et al., 2014].

Definition question patterns	Expected answer types
Who+be+<a question focus>+ ? من هو	من هي <الموضوع>+؟	*Interesting information about a person*
What+be+<a question focus>+? ما هو	ما هي<الموضوع>+؟	*Interesting information about an organization or a concept*

Table 8. An example of identifying a question focus.

Definition question	Lexical pattern	Named entity
Who is Steve Chen ?	*Who+be+<a question topic>+ ?*	*Steve Chen*
من هو ستيف تشين ؟	***من هو+ <الموضوع>+؟***	ستيف تشين

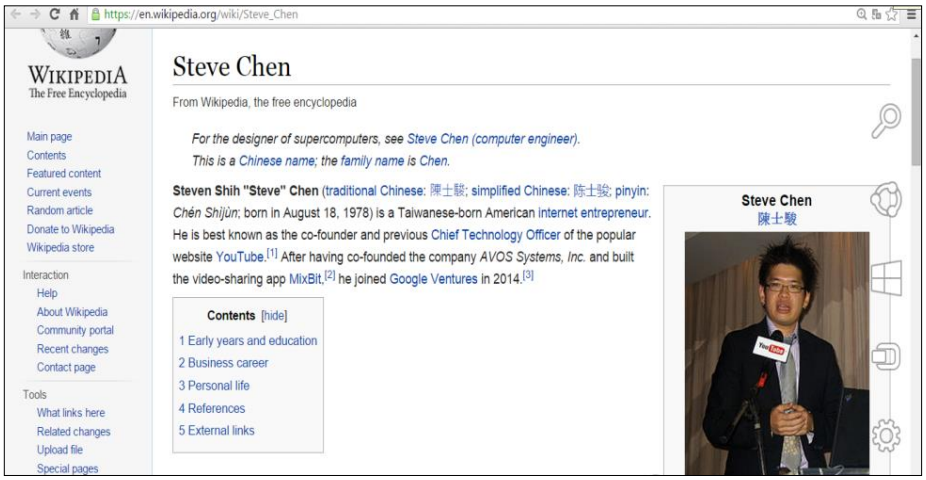


Fig. 8. An extract of a Wikipedia article entitled "Steve Chen".

Cooperative Answer Pattern Selection (sub-task 3.2). Two steps constitute this sub-task. The first consists in looking for the Wikipedia article infobox relative to the extracted question focus in the previous subtask. In case there is an infobox that which has a title identical to the question focus, the Wikipedia article infobox features are extracted and compared to the attributes of each infobox generated in task 1. The

second step consists in looking for the most suitable infobox class of the selected Wikipedia article infobox.

Fig.9. A Wikipedia article infobox entitled "Steve Chen" with its attributes.

A process of identifying the highest overlap rate between their attributes is lunched. The suitable infobox class selected is the one having the highest overlap rate of attributes. Its cooperative answer pattern is chosen as the adequate skeleton of the expected definition of the cooperative answer to generate. We take as an example the question cited in table 8 to show how to choose its suitable cooperative answer pattern. Figure 8 shows an extract of the Wikipedia article entitled "Steve Chen" (i.e. the identified question focus). To identify the adequate cooperative answer pattern, we compare the extracted infobox attributes (figure 9) to the attributes characterizing each generated infobox class. In this case, the cooperative answer pattern selected for the current question is the one assigned to the infobox class having as semantic category "entrepreneur" (see table 5).

Table 9. An example of a cooperative definition answer with its respective cooperative answer pattern.

<VALUE QUESTION FOCUS> + WAS +{BORN}+ IN + <VALUE> + . + HIS {RESIDENCE}+ IS +AT + <VALUE> + . + HIS +{ALMA MATER }+ IS +IN+ <VALUE> +.+ HIS + {OCCUPATION}+ IS/ARE + <VALUE> + , + {BOARD MEMBER OF} + <VALUE> +.+ HIS + {SPOUSE} + IS + <VALUE> +.+HIS + {WEBSITE} + IS + <VALUE> +.
Steve Chen was born in Taipei, Taiwan. His residence is in San Francisco, California, US. His alma mater is at the University of Illinois at Urbana-Champaign. His occupation is co-founder of Avos systems. His spouse is Park ji-hyun (Jamie Chen).
<المعلومة موضوع سؤال > {ولد/ت} + في + < المعلومة > +. + {تقيم/يقيم} + في+ <المعلومة> +.+ {متخرج/ة من} +<المعلومة> +.+ {يشغل منصب} + <المعلومة> +. + {شريك حياته/ها} + هو/هي + <المعلومة> +.+{ موقع الويب الخاص+}+هو+<المعلومة>+.
ستيف تشين ولد في تايبيه، تايوان. يقيم فيسان فرانسيسكو، كاليفورنيا، الولايات المتحدة. متخرج من جامعة إلينوي في أوربانا شامبين. يشغل منصب مؤسس مشارك ل"أفوس سيستم" . شريك حياته بارك جي-هيون (جيمي شين).

Answer Generation (subtask 3.3). It consists in generating definition answers by filling the blanks of the selected cooperative answer patterns by taking into consideration the correspondence between the cooperative answer pattern attributes and the extracted Wikipedia article infobox attributes. The generated answer is characterized by its cooperative form. For the question cited in table 8 "Who is Steve Chen?", the selected cooperative answer pattern is filled with the respective information extracted from the Wikipedia article infobox entitled "Steve Chen" (see figure 9). Each attribute of the cooperative answer pattern does not have its respective information is removed, while the other ones are kept. Table 9 shows the selected cooperative answer pattern and the generated cooperative answer in English and also with its translation in Arabic.

4 Answer Generation Module towards a Question Answering System

In order to evaluate the impact of the proposed approach for generating cooperative answers and to facilitate its integration in a definition question-answering system, we have implemented it in a module entitled cooperative answer generation. Two processing phases are required to realize this implementation of the proposed approach in a module: "An off-line processing phase" and "an on-line processing phase". The former includes common processing which is not associated with a given definition question (see figure 9) while the latter deals with treatments specific to a given defini-

tion question (see figure 11). As our approach can be applied to any language, we have decided to choose a low resource language, such as Arabic. This choice gives us more chance to deduce the limits and the performance in the difficult cases. The selected language is used for the resource, the questions and the answers.

The Off-line Processing Phase. This phase includes a series of treatments divided into two steps to exploit Wikipedia information. These steps are respectively "infobox class generation" and "cooperative answer pattern generation" (see figure 10).

Infobox Class Generation. This step consists in collecting Wikipedia articles, extracting the infoboxes, then grouping them in classes. The collection process consists in collecting all the Arabic Wikipedia articles from the website of Wikipedia[11]. The collection process is realized automatically through a tool entitled WikiPageDownload, developed by our research team. To a list of named entities, which exhibits the Wikipedia article titles in a given language, permits to download and save the content of each Wikipedia article associated with each named entity in the given list. For the Arabic language, 321454 Wikipedia articles are downloaded and saved. Figure 10 shows a distribution of these Wikipedia articles according to the existence of infoboxes, while figure 11 presents some examples of Wikipedia article titles.

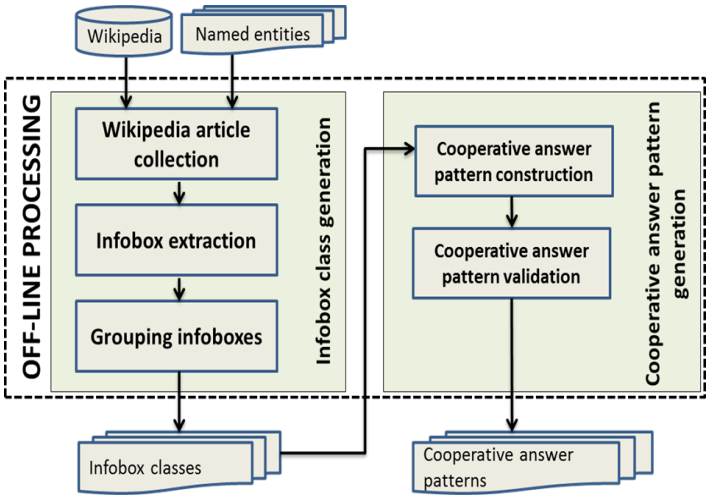


Fig. 10. The implementation process of the off-line phase.

The extraction of infoboxes process consists in extracting infobox features of 78760 downloaded Wikipedia article infoboxes (see Table 10). The process of grouping infoboxes consists in exploiting the downloaded infobox features by clustering together the infoboxes having highest overlap rate between their attributes. From

[11]https://sites.google.com/site/omartrigui/downloads

78760 Wikipedia article infoboxes, 189 infobox classes were generated; each infobox class contains nearly 417 infoboxes (see table 11).

Table 10. Arabic Wikipedia article distribution.

Wikipedia articles containing *infobo*xes	78760	**24,50%**
Wikipedia articles containing no infoboxes	242694	75,50%
Total Wikipedia articles in the Arabic language	321454	100%

Table 11. The average infoboxes per infobox class.

Number of infobox classes	189
Average infoboxes per infobox class	417

For each infobox class, a semantic category is attributed by a human expert presenting a hyperonymy of all their respective infobox titles. Tables 2, 3 and 4 present respectively three infobox classes respectively characterized by their attributes and their semantic categories, which are given hereafter: “entrepreneur”, “political party” and “sports team”.

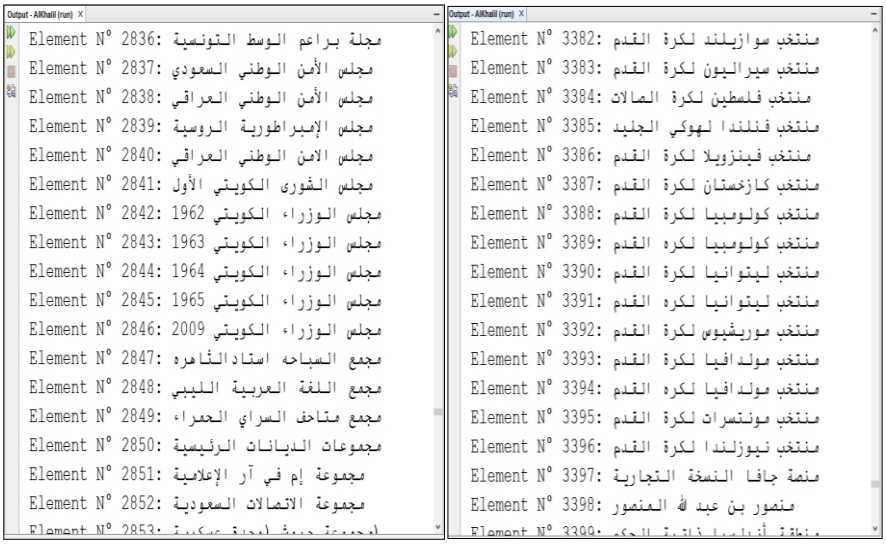


Fig. 11. An extract of the Wikipedia article titles list in the Arabic language.

Cooperative answer pattern generation. It consists in generating a cooperative answer pattern using infobox class attributes. For each infobox class, the respective attributes are taken in their order of appearance and followed by blanks and specific

punctuation marks (i.e. commas, full stops, etc.). This step enables to associate a cooperative answer pattern for each infobox class. For 189 infobox classes, we obtained 189 cooperative answer patterns.

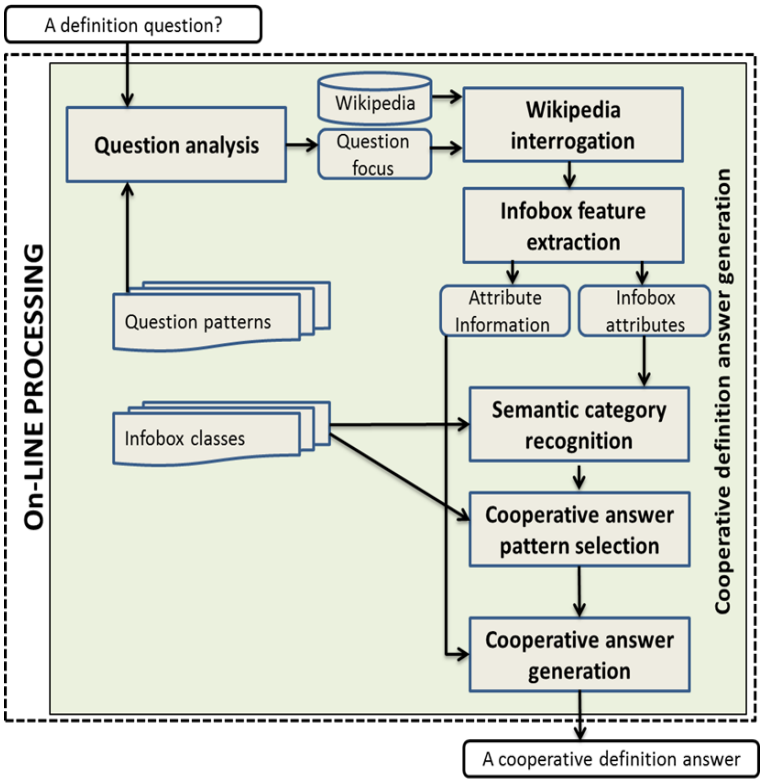


Fig. 12. The implementation process of the on-line phase.

The On-line Processing Phase. The on-line processing phase is based on the results of the off-line processing phase to generate cooperative definition answers. It is constituted by six processing steps: question analysis, Wikipedia interrogation, infobox feature extraction, semantic category recognition, cooperative answer pattern selection and cooperative answer generation (see figure 9). They are running for each given definition question to obtain the suitable cooperative definition answer. A cooperative definition answer does not exist in any documents but generated towards piece of information and a cooperative answer pattern.

5 Experiments

We will present the details of the experiments carried out to test the validity of the cooperative answer generation module and evaluate its impact. We have selected

DefArabicQA system (i.e. a definition question-answering system dealing with the Arabic language) to integrate this module and realize experiments with it [Trigui et al, 2010a]. DefArabicQA is a system based on both linguistic and frequency-based approaches. It uses a surface pattern technique to extract candidate answers and statistical features to rank them. It is based on Web search engines as knowledge bases [Trigui et al, 2010b]. Its architecture is illustrated in figure 13.

Experimental Data. Before illustrating the experimental results, we will describe the test data and the used performance measure. We have used a dataset comprising 300 definition questions in the Arabic language. The questioner is an adult, a native speaker of Arabic, and a reader of Arabic newspapers. Table 12 shows a part of these questions.

To measure the performance of the carried out experiments, the accuracy measure is used typically to evaluate the overall quality of a question answering system providing one potential answer for a given question. It is a number between 0 and 1, which indicates the probability of a question answering system to provide the correct answer on average. It is expressed as follows: *Accuracy = Number of correct answers / Number of questions*.

Evaluation Methodology. Three experiments are carried out. The first experiment is executed using the Google Web search engine (a baseline); the second one is realized using DefArabicQA system and the third experiment is performed using the DefArabicQA system extended by the cooperative answer generation module (see figure 14). All the experiment results are compared to a baseline.

The accuracy of the integrated module must be deduced. All the experiments are carried out with the same question dataset. We have to note that two assessors have evaluated the returned answers for each run. Both of them are Arabic native speakers and Arabic newspaper readers. To count the correct answers of these experiments with a fairer measure, we took the following hypothesis: for the first experiment (i.e. using the Google Web search engine), a question can be answered only if the first top snippet returned by the Google Web search engine containing at least one information nugget. For the second experiment with the DefArabicQA system, a question is annotated answered correctly only if its corresponding answer contains at least one information nugget without extraneous information. Then, for the third experiment with the DefArabicQA system extended with the cooperative answer generation module, a question can be answered correctly only if its answer contains a cooperative definition answer or at least one information nugget without extraneous information.

Experimental Results. We will now deal with the experimental results of the carried out experiments, which are presented in Table 13. Regarding the first experiment, 45% of the questions were answered by the Google Web search engine from the first top snippet but the search engine failed to return the correct answer from the first top snippet to the rest (55% of the total questions). This experiment obtained 0.45 as an accuracy measure. It was taken as a baseline to the other two experiments.

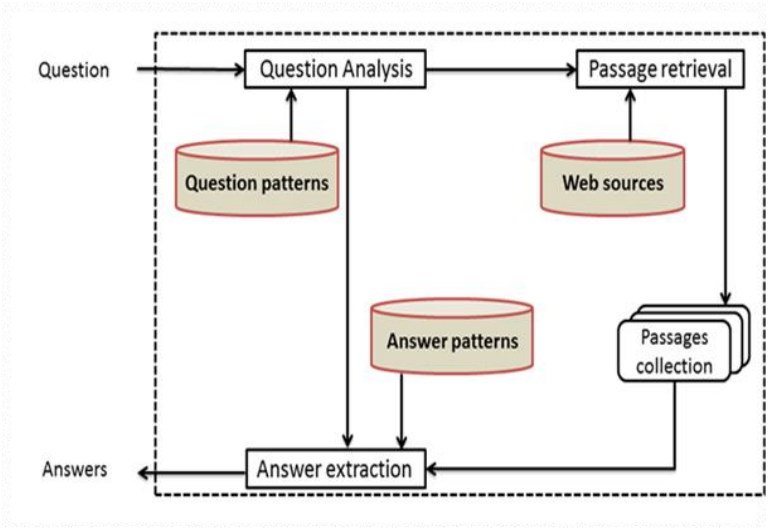


Fig. 13. DefArabicQA system architecture [Trigui et al, 2010b].

Table 12. A part of the definition questions of the test data.

Question N° 1 :	ما هو الاتحاد العام التونسي للشغل؟
	What is the Tunisian General Labor Union?
Question N° 2 :	ماهي الشركة السعودية للكهرباء؟
	What is the Saudi Electricity Company?
Question N° 3 :	ماهي شركة الزامل للاستثمار الصناعي؟
	What is the Zamit Industrial Investment Company?
Question N° 4 :	ماهي الشركة العامة للبريد والاتصالات السلكية واللاسلكية؟
	What is the General Post and Telecommunications Company?
Question N° 5 :	ماهي الشركة العربية للاستثمار؟
	What is the Arab Investment Company?
Question N° 6 :	ماهي الشركة العربية للعود؟
	What is the Arabian Oud Company?
Question N° 7 :	ماهي الشركة القابضة للنقل البحري و البري؟

	What is the Holding Company for Maritime and Land transport?
Question N° 8 :	ماهي الشركة العالمية للكتاب؟
	What is the World Book Publishing?

The results of the second experiment brought an overall improvement over the baseline. The DefArabicQA system succeeded in answering around 63%, but failed to answer around 37% of the questions. Accuracy is around 0.63, which it is over the baseline by 0.18.

In the third experiment, we tested whether the integration of the cooperative answer generation module can further improve the accuracy of the DefArabicQA system further. This experiment was carried out using DefArabicQA system extended by the cooperative answer generation module. A first part equal to 70% of the question set was answered; however, the remaining 30% of the questions were not answered in this experiment (see Table 13). The evaluation results obtained brought an improvement of 0.07 in the accuracy over the second experiment and 0.25 over the baseline.

Table 13. Experimental results.

	Answered questions (%)	Unanswered questions (%)
Google Web Search engine (baseline)	45	55
DefArabicQA system	63	37
Extended DefArabicQA system	70	30

5.1 Discussion

The results of the second experiment brought an overall improvement over the baseline. The DefArabicQA system succeeded in answering around 63%, which is important compared to the overall results released in TREC [Voorhees, 2003]. This fact explains why when we carried out the third experiment (i.e. using the DefArabicQA system with the cooperative answer generation module); we obtained an improvement of only 0.07% in the accuracy compared to the accuracy achieved in the second experiment (see figure 12 and Table 13). Indeed, a rise of the accuracy measure brought by the cooperative answer generation module, a significant improvement of the quality by more than 60% of the returned answers, is noticed. These answers have as a common feature their cooperative form; however, in the second experiment, they have been answered via the information nugget form.

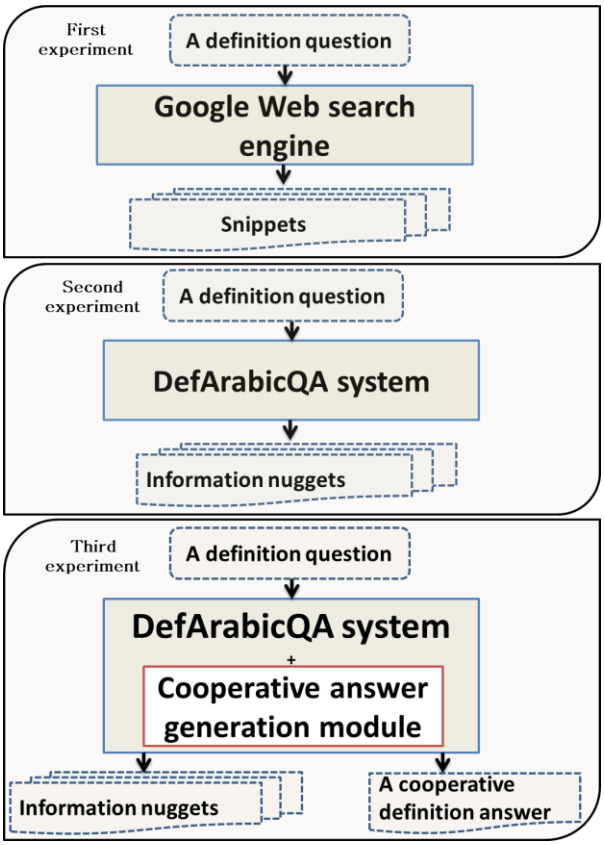


Fig. 14. The three carried out experiment architecture systems.

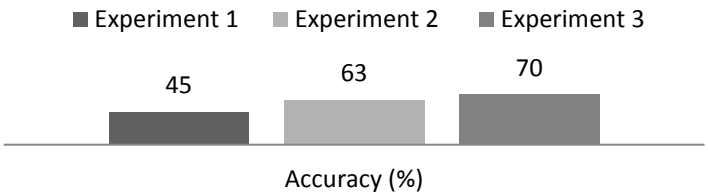


Fig. 15. Accuracy of the performed experiments.

We deduced from the percentage of questions answered by a cooperative answer that the cooperative answer generation module has succeeded in answering only from Wikipedia article infoboxes, up to 40% of the total test questions. This result is improved in spite of using only Wikipedia article infoboxes as a resource, knowing that the number of Arabic Wikipedia article infoboxes is very small compared to other languages. There are only around three hundred and twenty Wikipedia articles in Arabic compared to nearly four million Wikipedia articles in English. For the proposed module, the more Wikipedia articles there are, the more infobox classes we

have and the more successful we will be to generate a cooperative definition answer to a given definition question.

Thus, the fact that a portion of 60% of test questions which did not answered by a cooperative definition answer is caused mainly by the relative lack of Wikipedia articles in the Arabic language. In order to be more capable of dealing with this gap between languages, we try to look for information from Wikipedia article contents and not only from Wikipedia article infoboxes for certain languages. In general, the integration of the cooperative answer generation module in the DefArabicQA system has had a significant impact on the overall accuracy and on the quality of the returned answers. Therefore, it can be considered as a complementary module which has a positive influence on any definition question-answering system.

6 Conclusions and Future Work

We have proposed an approach for the generation of cooperative answers to definition questions. This approach is based on Wikipedia article infoboxes as a Web knowledge base. It is characterized by being language-independent and having the possibility to deal with open field questions. Its advantage is that it uses what is available on the Web to reach a cooperative definition answer, which should be consistent and informative, especially for a low resource language. Our experimental results show that the integration of the proposed approach in a definitional question-answering system dealing with the Arabic language has significantly outperformed the baseline which is based on Web search engines. In particular, we have shown that the Wikipedia article infoboxes can be used as a resource for generating definition cooperative answers. The limitations of the proposed approach are mainly related to the language adopted by the information resources.

As perspectives and in order to raise the effectiveness of this approach even if the specific Wikipedia articles do not contain infoboxes, we plan to exploit more Wikipedia article contents and test the validity of the approach in a multilingual context.

Acknowledgements. The work of the third author was partially funded by the Spanish Ministry of Economy, Industry and Competitiveness (MINECO) under the SomEMBED research project (TIN2015-71147-C2-1-P) and by the Generalitat Valenciana under the grant ALMAMATER (PrometeoII/2014/030).

References

1. Al-Bukhitan, S., Helmy, T., Al-Mulhem, M.: Semantic Annotation Tool for Annotating Arabic Web Documents. In: Proceedings of the 5th International Conference on Ambient Systems, Networks and Technologies (ANT-2014), vol. 32, pp. 429–436 (2014)
2. Al-Zoghbyaa, A. M., SharafEldin, A., Hamza, T. T.: Arabic Semantic Web Applications – A Survey. Emerging Technologies in Web Intelligence, vol. 5, no. 1, pp. 52–69 (2013)
3. Al-Zoghbyaa, A. M., Shaalan, K.: Conceptual Search for Arabic Web Content. In: Proceedings of the CICLing 2015, Part II, LNCS 9042, pp. 405–416 (2015)

4. Badawy, O., Shaheen, M., Hamadene, A.: ARQA High-Performance Arabic Question Answering System. In: Proceedings of Arabic Language Technology International Conference (ALTIC 2011), Alexandria, Egypt, pp. 129–136 (2011)
5. Benajiba, Y., Rosso, P., Abouenour, L., Trigui, O., Bouzoubaa, K., Belguith, H. L.: Question Answering. In: Natural Language Processing of Semitic Languages, pp. 335–370 (2014)
6. Benamara, F.: WebCoop: Un système de réponses coopératives sur le Web. Thèse de doctorat, Université Paul Sabatier (2004)
7. Benamara, F., Saint-Dizier, P.: Lexicalisation Strategies in Cooperative Question-Answering Systems. In: Proceedings of the 20th International Conference on Computational Linguistics (COLING), Geneva, Switzerland, pp. 345–352 (2004)
8. Beseiso, M., Ahmad, A. R., Ismail, R.: An Arabic Language Framework for Semantic Web. In: Proceedings of the Semantic Technology and Information Retrieval (STAIR), Putrajaya, pp. 7–11 (2011)
9. Bizer, C., Lehmann, J., Kobilarov, G., Auer, S., Becker, C., Cyganiak, R., Hellmann, S.: DBpedia - A Crystallization Point for the Web of Data. Web Semantics: Science, Services and Agents on the World Wide Web, Vol. 7, Issue 3, pp. 154–165 (2009)
10. Brzeski, A., Boinski, T.: Relation-based Wikipedia Search System for Factoid Questions Answering. International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 9, pp. 5601–5605 (2014)
11. Buscaldi, D., Rosso, P.: Mining knowledge from Wikipedia for the Question Answering task. In: Proceedings of the 5th International Conference on Language Resources and Evaluation (LREC 2006), Genova, Italy, pp. 727–730 (2006)
12. Corella, F., Lewison, K. P.: A Brief overview of cooperative Answering. http://pomcor.com/whitepapers/cooperative_responses.pdf (2009)
13. Cui, H., Kan, M. Y., Chua, T. S.: Soft Pattern Matching Models for Definitional Question Answering. ACM Transactions on Information Systems (TOIS), Vol. 25, Issue 2 (2007)
14. Dang, H. T., Kelly, D., Lin, J.: Overview of the TREC 2007 question answering track. In: Proceedings of the 16th Text REtrieval Conference (TREC 2007), Maryland, USA (2007)
15. Ezzeldin, A. M., Shaheen, M.: A Survey of Arabic question answering: challenges, tasks, approaches, tools, and future trends. In: Proceedings of the 13th International Arab Conference on Information Technology (ACIT 2012), Jordan, pp. 280–287 (2012)
16. Fareed, N., Mousa, H., Elsisi, A.: Syntactic Open Domain Arabic Question/Answering System for Factoid Questions. In: Proceedings of the 9th International Conference on Informatics and Systems (INFOS2014), Natural Language Processing and Knowledge Mining Track, pp. 1–9 (2014)
17. Hammo, B., Ableil, S., Lytinen, S., Evens, M.: Experimenting with a Question Answering system for the Arabic language. Computers and the Humanities, Vol. 38, Issue 4, pp. 397–415 (2004)
18. Hasanain, M., Elsayed, T., Magdy, W.: Identification of Answer-Seeking Questions in Arabic Microblogs. In: Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management (CIKM'2014), Shanghai, China, pp. 1839–1842 (2014)
19. Kurdi, H., Alkhaider, S., Alfaifi, N.: Development and Evaluation of a Web Based Question Answering System for Arabic Language. In: Proceedings of the Computer Science & Information Technology CS & IT-CSCP, pp. 187–202 (2014)
20. Khaled, S.: A Survey of Arabic Named Entity Recognition and Classification. Computational Linguistics, 40(2), pp. 469–510 (2014)

21. Kurdi, H., Alkhaider, S., Alfaifi, N.: Development and evaluation of a web based question answering system for Arabic language. International Journal on Natural Language Computing (IJNLC), Volume 3, Issue 2, pp. 11–32 (2014)
22. Ji, H., Grishman, R.: Knowledge Base Population: Successful Approaches and Challenges. In: Proceedings of the49th Annual Meeting of the Association for Computational Linguistics, USA, pp. 1148–1158 (2011)
23. Lopez, V., Uren, V., Sabou, M., Motta, E.: Is Question Answering fit for the Semantic Web?: a Survey. Semantic Web Journal, Volume 2, Issue 2, pp. 125–155 (2011)
24. Lupkowski, P., Leszczyńska-Jasion, D.: Generating cooperative question-responses by means of erotetic search scenarios. Journal of Logic and Logical Philosophy, Volume 24, pp. 61–78 (2014)
25. Shaheen, M., Magdy-Ezzeldin, A.: Arabic Question Answering: Systems, Resources, Tools, and Future Trends. Computer Engineering and Computer Science Arabian Journal for Science and Engineering, Volume 39, Issue 6, pp. 4541–4564 (2014)
26. Prager, J. M., Chu-Carroll, J., Czuba, K., Welty, C., Ittycheiach, A., Mahindru, R.: IBM's PIQUANT in TREC 2003. In: Proceedings of the 12th Text REtreival Conference (TREC 2003), Gathersburg, MD, pp. 283–292 (2003)
27. Pum-Mo, R., Myung-Gil, J., Hyun-Ki, K.: Open domain question answering using Wikipedia-based knowledge model. Information Processing & Management, Volume 50, Issue 5, pp. 683–692 (2014)
28. Surdeanu, M.: Overview of the TAC2013 Knowledge Base Population Evaluation: English Slot Filling and Temporal Slot Filling. In: Proceedings of Text Analysis Conference (TAC) (2013)
29. Trigui, O., Belguith, H. L., Rosso, P.: An automatic definition extraction in Arabic language. In: Proceedings of the 15th International Conference on Applications of Natural Language to Information Systems (NLDB 2010), Cardiff, UK, pp. 240–247 (2010)
30. Trigui, O., Belguith, H., Rosso, P.: DefArabicQA: Arabic Definition Question Answering System. In: Proceedings of the Workshop on Language Resources and Human Language Technologies for Semitic Languages, 7th LREC, Valletta, Malta (2010)
31. Trigui, O.: How to extract Arabic definitions from the Web? Arabic Definition Question Answering System. In: Proceedings of the 16th international conference on natural language processing and information systems (NLDB 2011), Alicante, Spain, pp. 318–323 (2011)
32. Voorhees, E. M.: Overview of the TREC 2003 Question Answering Track. In: Proceedings of the 12th Text REtrieval Conference (TREC 2003), pp. 54–68 (2003)
33. Voorhees, E. M.: Overview of the TREC 2004 Question Answering Track. In: Proceedings of the 13th Text REtreival Conference (TREC 2004), pp. 52–62 (2004)
34. Yang, M. C., Duan, N., Zhou, M., Rim, H. C.: Joint Relational Embeddings for Knowledge-based Question Answering. In: Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, pp. 645–650 (2014)
35. Yahya, A., Salhi, A.: Arabic Text Categorization Based on Arabic Wikipedia. ACM Transactions on Asian Language Information Processing (TALIP), Volume 13, Issue 1 (2014)

ISSN 1870-4069

Deformation and Residual Stress Based Multi-Objective Genetic Algorithm for Welding Sequence Optimization

Jesus Romero-Hdz[1], Gengis Toledo-Ramirez[1], Baidya Saha[2]

[1] Centro de Ingeniería y Desarrollo Industrial, Monterrey, Mexico

[2] Centro de Investigación en Matemáticas, Monterrey, Mexico

jaromero@cidesi.edu.mx, gengis.toledo@cidesi.edu.mx, baidya.saha@cimat.com

Abstract. Compared to deformation, residual stress has not been taken into account in the literature when it comes to welding process optimization. It also plays an important role to measure the weld quality. This paper reports the implementation of a multi-objective based Genetic Algorithm (GA) for welding sequence optimization, in which both structural deformation and residual stress are offered equal importance. The optimal weights between them are dynamically selected through optimizing a multi-objective fitness function in an iterative manner. A thermo-mechanical finite element analysis (FEA) was used to predict both deformation and residual stress. We chose the elitism selection approach to ensure that the three best individuals are copied over once into the next generation to facilitate convergence by preserving good candidates which can offer an optimal solution. We exploited a sequential string searching algorithm into single point crossover method to avoid the repetition of single beads into the sequence. We utilized a bit string mutation operator by changing the direction of the welding from one bead chosen randomly from the sequence. Welding simulation experiments were conducted on a typical widely used mounting bracket which has eight seams. Multi-objective based GA effectively reduces the computational complexity over exhaustive search with significant reduction of both structural deformation ($\sim$80%) and residual stress ($\sim$15%).

Keywords: Multi-objective optimization, genetic algorithm, residual stress.

1 Introduction

Fusion welding processes still very common in manufactutring, because they provide several advanteges in terms of cost, flexibility and design reliability. Gas Metal Arc Welding (GMAW) joins metals by simultaneously melting the base metal and adding a filler material to the joint to form a pool of molten material (the weld pool). Then, it cools down to form a joint that can be same or stronger than the base metal [9]. Welding is extensively used in a wide range of industries such as automotive, shipbuilding,

aerospace, construction, pipelines, nuclear, pressure vessels, heavy and earth-moving equipment [22,13].

The scope of this research is limited to GMAW process. Welding deformation and effective residual stress derive several negative impacts to the manufacturing process, adding additional cost in various ways, such as constraints in the design phase, extra operations, cost of quality and overall capital expenditure. Welding deformation and residual stress can be minimized by finding a suitable welding sequence. The industrial practice of welding sequence optimization is to select the best sequence by experience and sometimes running a simplified design of experiments which typically does not guarantee the optimal sequence [19]. Since conducting many real welding experiments is very expensive, welding deformation and effective residual stress are computed through a welding simulation software based on Finite Element Methods (FEM). Thermo-mechanical models can work under various welding conditions and geometric configurations. However, it is computationally very expensive and time consuming.

The best welding sequence can be achieved through a full factorial design procedure. For full factorial design, the total number of welding configurations can be computed by $N = n^r \times r!$, where n and r are the number of welding directions and beads (seams or segments) respectively. This number grows exponentially with the number of welding beads. For example, a complex weldment like an aero-engine assembly, it might have 52-64 weld segments [14]. Hence, the full factorial design is not feasible for industrial applications.

In this research, we implemented an iterative dynamic weight selection based multi-objective GA for welding sequence optimization. We make the following technical contributions in this paper:

- Multi-objective based GA effectively reduces the computational complexity over extensive search. In this research we have used eight weld seams and two welding directions. The number of welding configurations for exhaustive search is 10,321,920. In this experiment we achieved the optimal solution through GA after executing the welding simulation for 42 welding configurations. This is the minimum number of configurations necessary to find the optimal solution which was found based on the general Markov chain model of GA. The average execution time for each welding configuration using thermo-mechanical FEM approach is 30 minutes. Thus we saved significant amount of computational time.
- Literature reveals that both deformation and effective residual stress [6] measure the weld quality. Though, deformation was frequently used in the past studies, however, effective stress has been ignored as demonstrated in the Table 1. This paper combines both structural deformation and residual stress as a measure of weld quality and offers equal importance to both of them. The optimal weight was chosen through dynamic selection of weights in an iterative manner while optimizing the welding sequence optimization through multi-objective GA. We exploited a fitness function by the weighted linear combination of the inverse of the maximum structural deformation and effective residual stress.

- We expedited the convergence of the GA through the elitism selection approach and we copied the three best individuals into the next generation and preserved the best individuals which might lead to optimal solution.
- We tailored the single point crossover algorithm for the welding sequence optimization to avoid the repetition of single beads in the welding sequence by incorporating a sequential string searching algorithm into the single point crossover method.
- We implemented the bit string mutation algorithm by changing only the direction of the welding on one bead selected randomly from the sequence obtained by crossover algorithm instead of changing the bead itself to avoid the repetition of single bead in the welding sequence.

Experiments were conducted through the well-known simulation software Simufact Welding ®on a mounting bracket, which is widely used in telescopic jib [5] and automotive industries [26,12]. Experimental results demonstrate that best welding sequence can reduce significant amount of structural deformation ($\sim$80%), effective residual stress($\sim$15%) over the worst sequence.

The organization of the paper is as follows. Section 2 presents literature review. Section 3 discusses the thermal and mechanical analysis of FEM. Proposed dynamic weight selection based multi-objective GA for welding sequence optimization and its convergence analysis are presented in section 4. Results and discussions are demonstrated in section 5. Section 6 concludes this work. Relevant references are listed at the end of the paper.

2 Literature Review

Several authors have implemented GA for welding sequence optimization. A brief review is presented in Table 1. Furthermore, Table 2 shows a review of their validation methods.

Table 1. Literature review on implemented GA fitness function.

Author	Main functions				
	Trajectory time	Deformation	Residual stress	Temperature	Others
[27]	Yes	Yes	No	No	No
[17]	Yes	No	No	Yes	Robot joint movements
[20]	No	Yes	No	No	No
[23]	No	Yes	No	No	No
[4]	No	Yes	No	No	Stiffness and stress constraints
[13]	No	Yes	No	Yes	No
[15]	No	Yes	No	No	No

Among the studies mentioned in Table 1, Xie *et al.* [27] and Kim *et al.* [17] proposed multi-objective GA that are discussed below. Xie and Hsieh [27] have implemented

Table 2. Literature review on GA validation methods.

Author	Validation type				
	Specimen FEA	Specimen tryout	Real part FEA	Real part try-out	Others
[27]	No	No	Yes	No	No
[17]	No	No	No	No	Virtual trajectory Sim
[20]	No	No	Yes	No	No
[23]	Yes	No	No	No	No
[4]	No	No	Yes	No	No
[13]	Yes	Yes	Yes	No	No
[15]	Yes	No	No	No	No

GA for finding a combined clamping and welding sequence. A multi-objective fitness function is taken into account to minimize cycle time (gun travel path) and assembly deformation as shown in Equation 1. FEM was used to evaluate the fitness function on automotive parts by spot welding process.

$$\text{Min } F = w_1 \frac{D_i}{D_{0i}} + w_2 \frac{C}{C_0}, \quad i = 1,2,3...N, \tag{1}$$

where, w_1 and w_2 are weights that define the importance of each sub-function; D_i is the total deformation on every single node for the actual generation. D_{0i} is the total deformation on every single node for the initial generation; C is the cycle time for the actual generation and C_0 is the cycle time for the initial generation. Notice that $\frac{D_i}{D_{0i}}$ and $\frac{C}{C_0}$ are considered as normalized functions because the units of deformation and cycle time are different.

Kim *et al.* [17] have implemented GA using a multi-criteria fitness function (Equation 2). This function includes the minimization of gun travel time, avoidance of thermal distortion and smooth robot joint movement. The criteria considered here are Euclidian distance between weld seams, a 30 mm distance considered as heat affected zone and total change of the robot joints respectively. This algorithm is suitable for different arc welding operations such as multi weld lines: singlepass or multipass:

$$\text{Min } F = \text{Min}(w_1 g_1 + w_2 g_2), \tag{2}$$

where: w_1 and w_2 are weights. The sub-function that involves gun travel time and distortion criteria g_1 is defined by

$$g_1 = \sum_{a_{ij} \in T} x_{ij}, \tag{3}$$

where: T is a trajectory,

$$x_{ij} = \begin{Bmatrix} c_{ij} \text{ if } a_{ij} \notin h_{ij} \\ c_{ij} + M_1 \text{ if } a_{ij} \in h_{ij} \end{Bmatrix}, \tag{4}$$

$$c_{ij} = \begin{Bmatrix} l_{ij} \text{ if } a_{ij} \in W \\ l_{ij} + M_2 \text{ if } a_{ij} \notin W \end{Bmatrix}, \tag{5}$$

where: h_{ij} is the heat affected zone for each weld seam a_{ij} in W; l_{ij} is the arc length a_{ij}; A is a set of arcs a_{ij} from each node $i \in N$ to each node $j \in N$; N is a finite set of nodes in the seam w. W is a set of arcs that represents a weld seam $W \subseteq A$. For the sub-function that involves the smooth robot joint movements g_2 is defined by:

$$g_2 = \sum_{a_{ij} \in T} \sum_{k \in J} \theta_{ijk}, \tag{6}$$

where: θ_{ijk} is the angle of change for a joint k from one node i to other node j from the set a_{ij}. J is a set of robot joints. The penalty terms M_1 and M_2 are sufficiently large numbers. M_1 ensures that only seams out of the heat affected zone criteria (30 mm) will be selected. M_2 ensures that only valid segments are selected and all of them will be traveled.

From Table 1, it was shown that past studies have considered deformation and some of them exploited trajectory time of the robot, temperature, robot joint movements, stiffness and stress constraints as the measure of the weld quality. However, residual stress, which is also an important measure for the weld quality [6] has been ignored in the GA studies utilized for weld sequence optimization. In this research we exploited both deformation and effective residual stress as the measure of the weld quality and exploited both of them equally in the fitness function using iterative dynamic weight selection based multi-objective GA that are discussed in the subsequent sections.

3 Welding Simulation Framework

In order to present our approach we overview the welding simulation framework. This is important because the fitness function is computed using FEA.

3.1 Thermal Analysis

Weld process modeling (WPM) is a very complex task. The physics of heat generation has as a fundamental principle the law of conservation of energy. The heat equation can be written in the following form [9] (**Conservation of Energy or Heat Equation**):

$$\dot{h} = \nabla . q + Q, \tag{7a}$$

$$q = -\kappa \nabla T, \tag{7b}$$

$$dh = \rho c_p dT, \tag{7c}$$

where h, q, Q, T and ∇T represent enthalpy, thermal flux, power density function, temperature and temperature gradient respectively. κ and c_p represent thermal conductivity tensor and specific heat respectively and both are temperature dependent material properties. Temperature history in every node is computed by the transient heat conduction equation 7a, where the change in enthalpy $\dot{h} = \rho C_p \frac{\partial T}{\partial t}$, where: ρ is the material density

(g/mm^3), C_p specific heat capacity $(J/(g°C))$, T is the current temperature $(°C)$, q flux heat vector (W/mm^2), Q is the internal heat generation rate (W/mm^3), t is the time (s), ∇ is the spatial gradient operator x,y,z $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z})$. The heat flux is the amount of energy that flows through a particular surface per unit area per unit time. This is defined by the differential form of Fourier's Law of thermal conduction equation described in 7b. Here κ is the temperature-dependent thermal conductivity matrix $(J/mms°C)$ and ∇T is the temperature gradient.

Typically, the complexity of the heat generation physics in the weld puddle is simplified by using a heat input model or well known as welding simulation models. The classical approach in Computational Welding Mechanics (CWM) is to ignore fluid flow and use a heat input model where heat distribution is prescribed. The given heat input replaces the details of the heat generation process and focus on larger scales. Moreover, the modeling of fluid flow and pertaining convective heat transfer may be integrated with a CWM model.

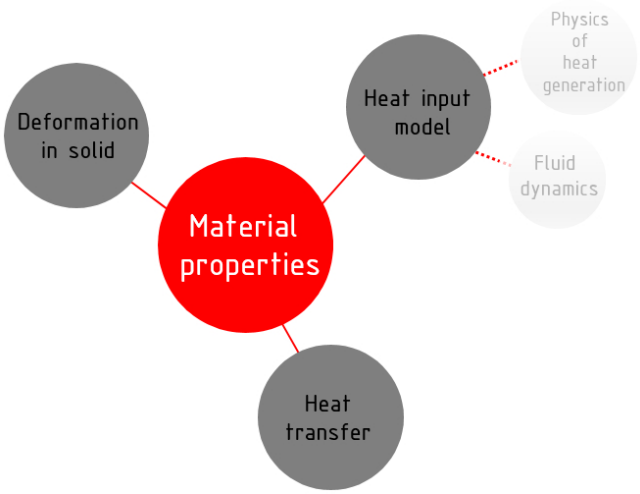


Fig. 1. Welding simulation.

The most common used model for fusion welding processes is the well-known Goldak double ellipsoidal heat distribution. This heat input model combines two ellipsoidal heat sources to achieve the expected steeper temperature gradient in front of the heat source and a less steep gradient at the trailing edge of molten pool. This two heat sources are defined by Front heat distribution:

$$Q(x',y',z',t) = \frac{6\sqrt{3}f_f Q_w}{\pi\sqrt{\pi}abc_f} e^{(\frac{-3x'^2}{a^2})} e^{(\frac{-3y'^2}{b^2})} e^{(\frac{-3z'^2}{c_f{}^2})}. \quad (8)$$

Rear heat distribution:

$$Q(x',y',z',t) = \frac{6\sqrt{3}f_r Q_w}{\pi\sqrt{\pi}abc_r} e^{(\frac{-3x'^2}{a^2})} e^{(\frac{-3y'^2}{b^2})} e^{(\frac{-3z'^2}{c_r{}^2})}, \tag{9}$$

where: f_f is the fraction factor of heat deposited in the front part, f_r is the fraction factor of heat deposited in the rear part. Those factors must satisfy the relation $f_f + f_r = 2$. a is the width, b is depth, c_r is the length of the rear ellipsoid y c_f is the length of the front ellipsoid.

These parameters are physically related to the shape of the weld puddle Figure 2. Width and depth are commonly taken from the cross section, the authors recommend to use a half of parameter a for the front fraction and two times a for the rear fraction. For a linear trajectory along axis z, is defined by z':

$$z' = z + v(\tau - t), \tag{10}$$

where z actual coordinate z, v is travel speed, τ is a delay factor and t is the time. The heat available from the heat source is defined by:

$$Q_w = \eta IE, \tag{11}$$

where η heat source efficiency, I is the current (A) , E is the voltage (V).

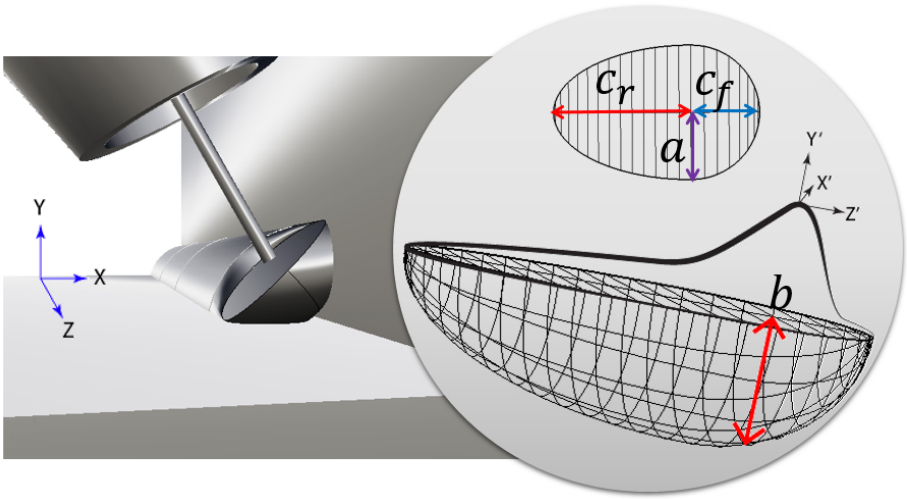


Fig. 2. Goldak double ellipsoidal model.

Thus the heat input model in CWM must be calibrated with respect to experiments or obtained from WPM models. Therefore, the classical CWM models have some limitations in their predictive power when used to solve different engineering problems. For example, they cannot prescribe what penetration a given welding procedure will

give. The appropriate procedure to determine the heat input model is therefore important in CWM [21].

The FEM software solves this time dependent system of partial differential equations on a domain defined by a FEM mesh. The domain is dynamic in that it changes with each time step as filler metal is added to the weld pass. The initial condition is often assumed to be the ambient temperature but the domain can be initialized to any initial temperature field. The heating effect of the arc is often modeled by a double ellipsoid power density distribution that approximates the weld pool as measured from macrographs of the cross-section of several weld passes. A convection boundary condition $q = h(T - T_{amb})$ with convection coefficient h and ambient temperature T_m usually is applied to external surfaces. The FEM formulation of the heat equation leads to a set of ordinary differential equations that are integrated in time using a backward Euler integration scheme.

3.2 Mechanical Analysis

The temperature history from the thermal analysis was used as a series of loads in the structural analysis. In this phase, the temperature history from the thermal cycle of each node is taken as an input and it is used as a node load with temperature dependent material properties. The solid model mesh used for the mechanical analysis was also used for the thermal analysis where each increment of weld deposition corresponded to one load step. Because phase transformation has an insignificant effect on the welding residual stress and distortion, the total strain ε^{total} (assuming negligible contribution from solid state phase transformation) can be decomposed into three components as follows: $\varepsilon^{total} = \varepsilon^{e} + \varepsilon^{p} + \varepsilon^{th}$, where $\varepsilon^{e}, \varepsilon^{p}$, and ε^{th} represent elastic, plastic and thermal strain respectively. In the welding process, changes in stress caused by deformation are assumed to travel slowly compared to the speed of sound. So, at any instant, an observed group of material particles is approximately in static equilibrium, i.e., inertial forces are neglected.

In rate independent plasticity, viscosity is zero and viscous forces are zero. In either the Lagrangian or the Eulerian reference frame, the partial differential equation of equilibrium is, at any moment is given by the conservation of momentum equation that is mentioned below [8] (**Conservation of Momentum Equation**):

$$\begin{aligned} \nabla.\sigma + f &= 0, \\ \sigma &= D\varepsilon, \\ \varepsilon &= (\nabla u + (\nabla u)^T + (\nabla u)^T \nabla u)/2, \end{aligned} \tag{12}$$

where $\nabla, \sigma, f, D, \varepsilon$ and u represent partial differential, cauchy strss, total body force, temperature dependent material property (elastic matrix relevant to the modulus of elasticity and Poisson's ratio), the Green-Lagrange strain and displacement vector respectively. ∇u represents the displacement gradient.

The mechanical model is based on the solution of three partial differential equations of force equilibrium illustrated in Equation 12. In the FEM formulation, Equation 12 is transformed and integrated over the physical domain, or a reference domain with a

unique mapping to the physical domain [9]. The simulation software solves this partial differential equation for a viscothermo-elasto-plastic stress-strain relationship. The initial state often is assumed to be stress free. Dirichlet boundary conditions constrain the rigid body modes. The system is solved using a time marching scheme with time step lengths of approximately 0.1 second during welding and 5 second during cooling phase.

4 Proposed Genetic Algorithm Based Welding Sequence Optimization framework

Genetic Algorithms emulate natural selection of a set of individuals in order to search the best solution to a problem [10]. The genetic configuration of each individual is a possible solution. GA starts with an initial population and those are submitted to an evolutionary process in such way that the best adapted individuals will continue to reproduce among them and over several generations the best adapted stands out. We tailor the GA for the welding sequence optimization: selection, cross-over, and mutation to avoid the repetition of single bead that is discussed below.

4.1 String Representation of Welding Sequence

Being Q the welding application and S a set of all possible sequences of Q, each sequence $s \in S$ represents a possible sequence which minimizes the overall structure deformation and residual stress. Each sequence has N weld seams, here called genes $s = \{x^1, x^2, x^3, \ldots x^N\}$, these are a combination of real numbers $\forall n = 1, 2, 3, \ldots N$. In this approach every seam can be welded in two directions and it is represented by a positive sign $if \circlearrowleft or \uparrow or \leftarrow$ or negative sign $if \circlearrowright or \downarrow or \rightarrow$.

4.2 Initialization of welding sequence

The algorithm starts with an initial population $P = \{s_j\}$, where elements of the set of sequences are called "individuals" $j = 1, 2, 3, \ldots J$. Their genes are generated randomly and special considerations taken in order to avoid repeated seam in the same welding sequence.

4.3 Deformation Based Fitness Value

Within the scope of natural selection, the individual eligibility is regarded as the degree of adaptability. In this paper we have implemented a multi-objective fitness function that takes into account deformation and residual stress and returns a real number (weighted linear combination of maximum deformation and residual stress of the structure) $f\left(s_{j=1}^{J}\right) \Rightarrow \mathbb{R}$ that measures the adaptability of each sequence:

$$F(s_j) = \sum_{i=1}^{I} w_i f_i, \tag{13}$$

where w_i is the weight that defines the importance of each sub-function. These weights are computed dynamically in each iteration t with the equation 14 [7]:

$$w_i(t) = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{I} \left| f_j(x_{t-1}) \right|}{(I-1) \times \sum_{j=1}^{I} \left| f_j(x_{t-1}) \right|}, \tag{14}$$

where I is the number of sub-functions, $i = 1...I$, x_{t-1} is the best individual among solutions of the population in the previous generation P_{t-1}. $w_i(t)$ is the dynamic weight satisfying the following conditions,

$$0 \leq w_i(t) \leq 1 \text{ and } \sum_{i=1}^{I} w_i(t) = 1, \tag{15}$$

where t represents the iteration step of the GA algorithm. f_1 is a sub-function that takes into account the final deformation on the structure and it is computed by FEA. Final deformation is defined by the equation 18:

$$f(s_j) = 1/(Max(D_i) + \varepsilon), \tag{16}$$

where D_i is the total deformation on every node defined by

$$D_i = \sqrt{{d_{x_i}}^2 + {d_{y_i}}^2 + {d_{z_i}}^2}, \tag{17}$$

$$i = 1, 2, 3...N,$$

d_{x_i}, d_{y_i}, and d_{z_i} are the deformation of node i along x, y, and z axis respectively. ε is a very small number which was used to offer continuity to the fitness function when the value of the maximum deformation is zero.

f_2 is a sub-function that takes into account effective stress. It is also computed by FEA and it is defined by the following equation .

$$f(s_j) = 1/(Max(E_i) + \varepsilon), \tag{18}$$

$$E_i = \frac{\sqrt{(\sigma_{1i} - \sigma_{2i})^2 + (\sigma_{2i} - \sigma_{3i})^2 + (\sigma_{3i} - \sigma_{1i})^2}}{\sqrt{2}}, \tag{19}$$

$$i = 1, 2, 3...N,$$

where σ_1, σ_2, σ_3 are the maximum, intermediate and minimum principal stresses. ε is a very small number which was used to offer continuity to the fitness function when the value of the maximum effective stress is zero.

4.4 Welding Sequence Selection Algorithm

Selection is an important sub-routine where individuals are chosen from the actual population for later procreation. Good selection algorithm expedites the convergence of the welding sequence. As a selection procedure, we first implemented a truncation procedure where the population is sorted by ascending fitness values, then a proportion μ of the individuals are taken based on fitness value. The proportion μ is computed by the fraction of the individual fitness value to the sum of the fitness values of all the samples as shown in Fig. 3.

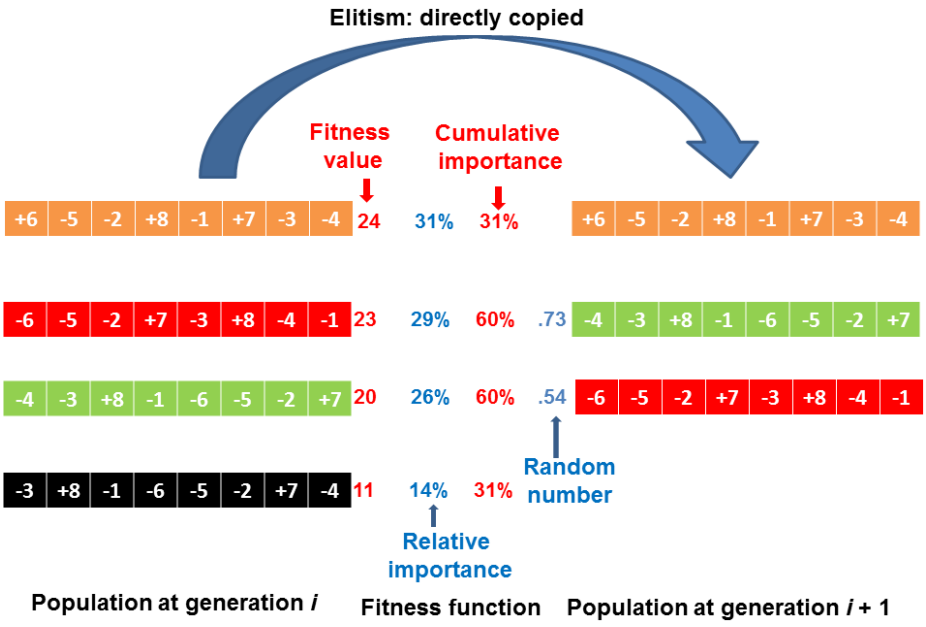


Fig. 3. Selection procedure with elitism function.

4.5 Crossover for Generating New Welding Sequences

Crossover is analogous to reproduction, new individuals are created from the selected parents. Each couple of selected individuals s_1 and s_2 exchange their genes and make two new individuals, $s'_1 = s_1 \times s_2$ and $s'_2 = s_2 \times s_1$. Several methods for crossover are reported in literature such as arithmetic, heuristic, single or multi-point, uniform, cycle, partially mapped and order [16,11]. In this paper we implemented a single point crossover as demonstrated in Fig. 4 where a random number defines the cut point $a \in [1,N]$. Later, the descendants are defined by equations 20 y 21 respectively.

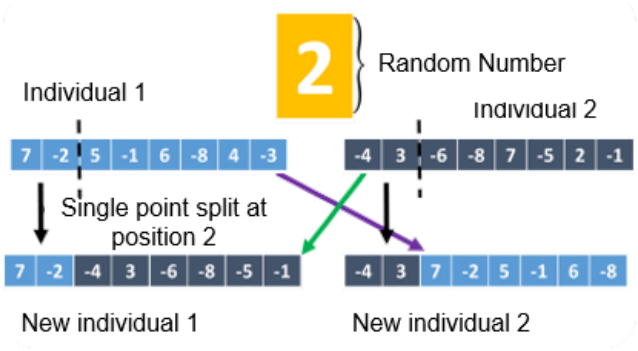


Fig. 4. Single point crossover operator.

To avoid the repetition of the weld seam in the same welding sequence during crossover we implement a repeated string validation algorithm, the pseudo-code of which is illustrated below.

function REPEATED STRING VALIDATION
  random number $a \in [1, N-1]$;
  $s'_1 = \{x_1^1, \ldots x_1^a\}$;
  $s'_2 = \{x_2^{a+1}, \ldots x_2^N\}$;
  for $i = 1 : N$ **do**
    if $\Pi(\sqrt{s'_1 . * s'_1}) \neq \sqrt{s_2(i)' . * s_2(i)'}$ **then**
      $s'_1 = \{[s'_1] \cup s_2(i)'\}$;
    end if
    if $\Pi(\sqrt{s'_2 . * s'_2}) \neq \sqrt{s_1(i)' . * s_1(i)'}$ **then**
      $s'_2 = \{[s'_2] \cup s_1(i)'\}$;
    end if
  end for
end function

$$s'_1 = \left\{[x_1^1, \ldots x_1^a], [x_2^{a+1}, \ldots x_2^N]\right\}, \tag{20}$$

$$s'_2 = \left\{[x_2^1, \ldots x_2^a], [x_1^{a+1}, \ldots x_1^N]\right\}. \tag{21}$$

4.6 Mutation for Generating New Welding Sequences

Mutation alters one or more individual genes from its actual configuration. It occurs during evolution in a low incidence according to a defined mutation probability. Some of the operators found in literature are bit string, delta, invert and swap [24,1,18]. Here we have used a bit string operator in order to change the direction of welding only rather than the welding seam itself as shown in Fig. 5 to avoid the repetition of the weld seam.

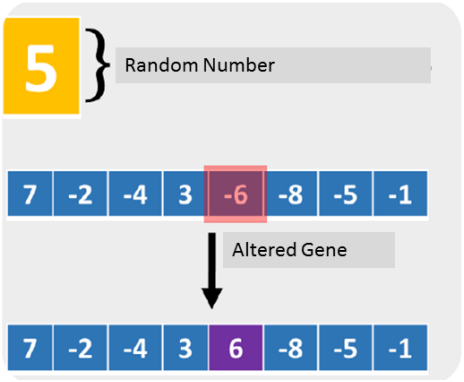


Fig. 5. Bit string mutation operator.

4.7 Elitism Based Welding Sequence Selection Algorithm

The *Elitism* function is a practical variant that ensures that the best individual in the actual population $s_{best} \in P_t$ and current generation t to carry over to the next generation P_{t+1} as shown in Fig. 3. Elitist based selection algorithm guarantees that the convergence obtained by the GA will follow monotone decreasing behaviour over generations, $[s_{best} \in P_t] \rightarrow P_{t+1}$.

4.8 Pseudo-code and Flowchart of the Proposed Iterative Genetic Algorithm for the Welding Sequence Optimization

The following algorithm is a repetitive process where the population is going to be changing over the generations $P_t = (s_1(t),\ s_2(t), ...s_J(t)) \in S$. The pseudo-code for the proposed GA based welding sequence optimization is given below.

function GA($Min\ D : Q$)
Input: $P_0 = (s_1(t),\ s_2(t), ...,\ s_J(t)) \in S$
Output: s_{best}, the best sequence that shows the minimum value of the weighted linear combination of deformation and residual stress.

$t \leftarrow 0$;
Initialize $P_t \in S$;
We assign arbitrary positive real numbers to $w_i(0), i = 1, ..., I$, satisfying the conditions mentioned in equation 15.
while !terminating condition **do**
$t++$;
Compute the fitness function $F(s_j) = \sum_{i=1}^{I} w_i(t-1) f_i(t-1);\ j = 1, 2, ..., J$ $\forall s_j \in P_{t-1}$
Select P_t from P_{t-1} based on the relative importance of the value of the individual fitness function $F(s_j)$. ; /* Priority given to the welding sequences based on weighted linear function of less deformation and residual stress */

Crossover $P_t \leftarrow P_t$; /* String searching based single point crossover */
Mutation $P_t \leftarrow P_t$; /* Change the direction of the welding of one seam */
Evaluate $F(s_j), j = 1,2,...,J$;
elitism $P_t \leftarrow s_{best}$ from P_t; /* Elitism based selection approach */
Compute $f_i(s_{best}),\ i = 1,,2,..,I$.

If $\sum_{i=1}^{I} |f_i(s_{best})| \neq 0$ then compute, $w_i(t) = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{I} |f_j(s_{best})|}{(I-1)\times\sum_{j=1}^{I} |f_j(s_{best})|}, i = 1,2,...,I;$

end while
return s_{best} from P_t.
end function

Fig. 6 describes the flowchart of the GA based welding sequence optimization approach.

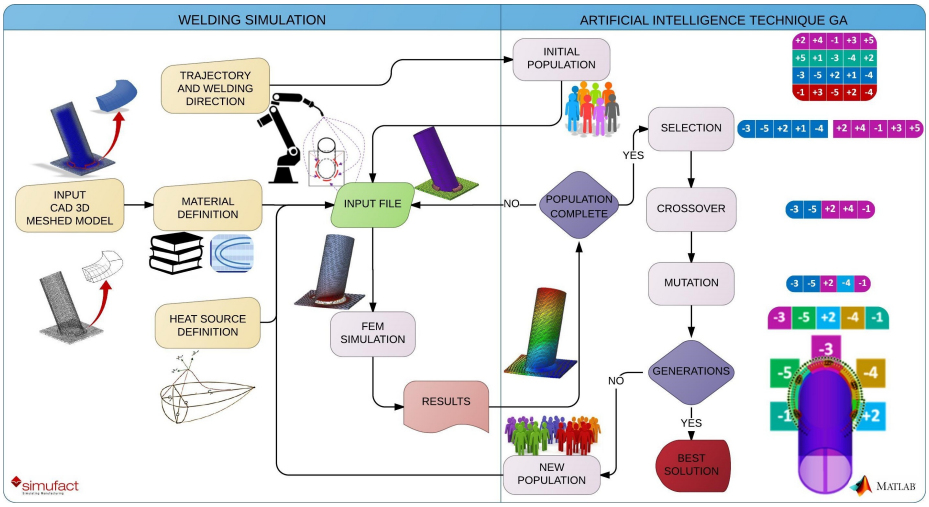


Fig. 6. GA based welding sequence optimization approach.

4.9 Convergence Analysis of the Genetic Algorithm

Aytug and Koehler [2,3] showed that for a general Markov chain model of genetic algorithm with elitism, an upper bound for the number of iterations t required to generate a population S^+ which consists entirely of minimal solutions has been generated with probability $\alpha \in (0,1)$, is given by:

$$t \geq \lceil \frac{ln(1-\alpha)}{nln(1-min\{\mu^l,(1-\mu)^l\})} \rceil, \tag{22}$$

where, l is the length of the chains that represent the individual, n is the population size and $\mu \in (0,1)$ is the mutation rate. $\lceil x \rceil$ is the smallest integer greater than or equal to x. Studniarski [25] showed that for multi-objective optimization, the (possibly unknown) number m of these solutions is bounded from below by some known positive integer $\bar{m}$. Suppose also that there exists a number $\beta \in (0, 1/\bar{m})$, an upper bound for the number of iterations t is given by:

$$t \geq \lceil \frac{ln(1-\alpha)}{ln(1-(\bar{m}\beta)^l)} \rceil. \tag{23}$$

If no non-trivial lower bound $\bar{m}$ is known, we may always use $\bar{m} = 1$.

5 Experimental Results

This section consist of the following subsections, fist we describe the study case. Second, we introduce the parameters we have used in terms of GA configuration. Third, convergence analysis of multi-objective GA is given and last, we present the effects of welding sequence optimization over the quality variables we targeted before.

5.1 Study Case

We chose a study case of welding a mounting bracket shown in Fig. 7 and 8 which is typically used in telescopic jib [5], automotive industries [26,12]. We conducted a simulation experiment of GMAW using popular Simufact® welding software. For details about the software, please see [13]. We implemented a multi-objective GA algorithm for choosing the best welding sequence having minimum weighted linear combination of structural deformation and effective residual stress and we demonstrated the effects of welding sequence on the weld quality (structural deformation and effective stress) by analyzing the structural deformation and residual stress caused by welding of the four sequences (best, second best, worst and second worst found by GA). Fig. 7 shows geometries of different mounting brackets that can be found frequently in heavy equipment, vehicles, ships, and Fig. 8 illustrates the engineering drawing with all specifications of the mounting bracket used in this experiment.

Fig. 7. Different mounting brackets in the market.

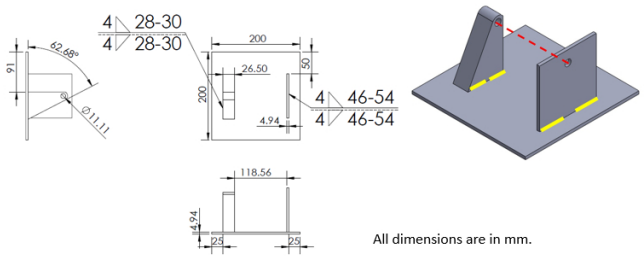


Fig. 8. Study case with 8 seams.

5.2 Parameters Used for this Study

Table 3 shows the GA parameters used in the simulation experiment. We considered 7 generations to converge the GA, initial population size as 6, crossover probability as 50%. We copy three best candidates of the current generation to the next generation using elitism based selection mechanism. We implemented single point cross-over method for new sample reproduction. We also implemented single bit string mutation operator and changed the welding direction of a randomly selected welding seam instead of welding seam itself to avoid the repetition of the welding seam in the sequence.

Table 3. GA parameters.

Parameter	Value
Initial population size	6
Generations	7
Elitism candidates	3
Crossover %	50%
Mutation operator	bit string
Crossover operator	single point
Qty of seams	8
Possible welding directions	2

5.3 Convergence Analysis of Multi-objective GA

We carried out a multi-objective GA experiment for seven generations and conducted the convergence analysis. Fig. 9, 10 11 illustrate the behaviors of the four individuals (best, second best, worst and second worst) in terms of deformation, effective stress and multi-objective criteria as mentioned in equation 13. Elitism based selection method expedites the convergence of the GA. Figures show the monotonically decreasing values of the deformation over seven generations.

The best, second best, worst and second worst sequences are (+6, -5, -2, +8, -1, +7, -3, -4),(-6, -5, -2, +7, -3, +8, -4, -1), (-4, -3, +8, -1, -6, -5, -2, +7) and (-3, +8, -1, -6, -5, -2,

+7, -4) respectively as can be seen in Figure 12. Their maximum structural deformation and effective residual stress values are $0.65mm.$, $0.82mm.$, $2.0mm.$, and $2.08mm.$ and $427.68MPa$, $424.40MPa$, $425.20MPa$, and $425.96MPa$ respectively as shown in Figures 13 and 14.

To compute the minimum number of iterations necessary to ensure finding an optimal solution for GA with a prescribed probability $\alpha = 0.98, \bar{m} = 1$, $\beta = 0.9$, number of bits required to represent an individual $l = 8$ in equation 23, we get $t \geq \lceil 6.95 \rceil = 7$. We conduct the multi-objective GA upto seven iterations since the computational complexity of the FEM based thermo-mechanical welding simulation approach is computationally very expensive.

Computation time. The average computation time for every individual of the GA using athermo-mechanical FEA approach using Simufact Welding ®simulation software was 30 minutes. This time depends on the convergence of the thermal analysis.

Fig. 9. Deformation.

Fig. 10. Effective Stress.

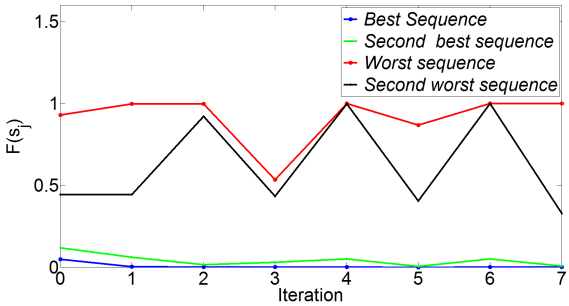


Fig. 11. Multi-Objective.

Figure 13 shows deformation distribution patterns for best, second best, second worst and worst sequences. It is clear that in worst and second worst sequences the maximum deformation value is greater than the best and second best sequences. Moreover, the distribution pattern is also better in the best ones, because the area near to the holes in the vertical pieces is less affected to pass a shaft or a bar through those holes.

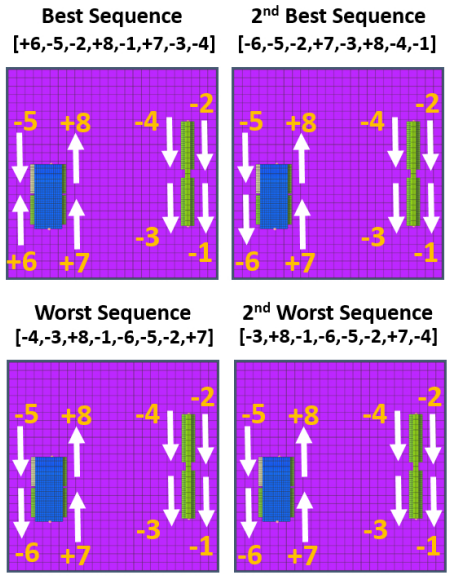


Fig. 12. Best, second best, worst and second worst sequences and their corresponding maximum structural deformation and effective stress.

Figure 14 shows effective residual stress distribution patterns for best, second best, second worst and worst sequences. The patterns are quite similar at the first looking,

however there are differences in their frequency. The reduction in magnitude is lower than deformation. It is clear that effective residual stress affects the area close to the welding seam.

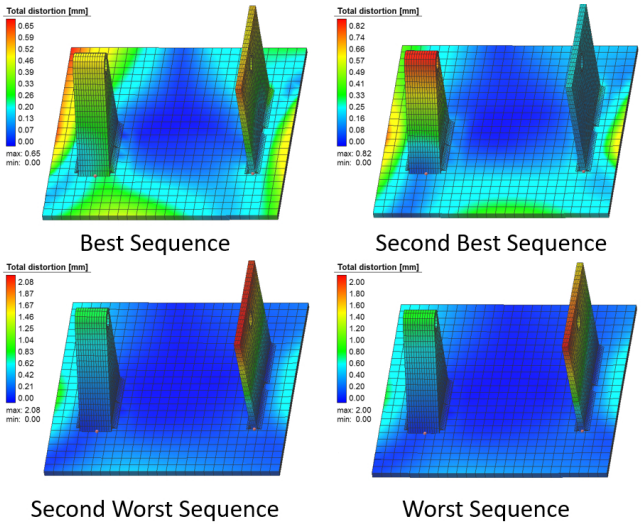


Fig. 13. Deformation patterns for best,second best, second worst and worst sequences.

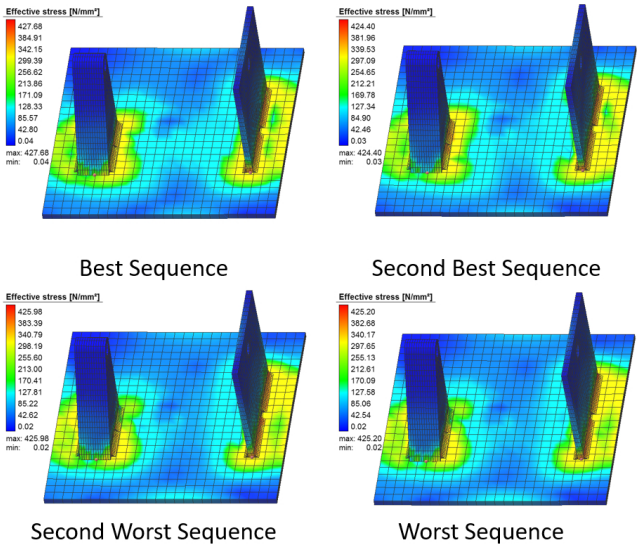


Fig. 14. Effective residual stress patterns for best,second best, second worst and worst sequences.

Fig. 15 shows the values of the two terms, $w_1 f_1$ and $w_2 f_2$ of the multi-objective function over seven iterations of GA. This graph shows that the dynamic weighted method ensures an equitable treatment of each objective [7]. In this experiment, we always have $|w_1 f_1 - w_2 f_2| \leq 0.07$ as shown in Fig. 15. Fig. 16 illustrates the Pareto front at iteration 1, 4 and 7.

Results of the Pareto front shown in Fig. 16 demonstrate that both structural deformation and effective stress cannot be simultaneously reduced, decreasing the value of one increases the value of the other and vice versa. In this study, utilizing iterative dynamic weight selection based multi-objective GA, we find the optimal weight between deformation and effective residual stress through fair treatment of both of them.

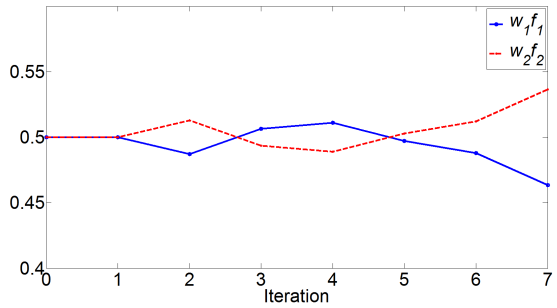


Fig. 15. $w_1 f_1$ and $w_2 f_2$ over dynamic weighted based multi-objective GA.

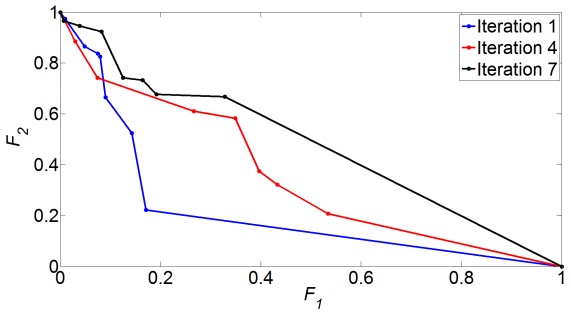


Fig. 16. Pareto Front.

5.4 Effects of Welding Sequence on Welding Process Optimization

Fig. 17, 18 and 19 demonstrate the normalized frequency of the deformation, effective residual stress and Multi-objective function (Equation 13) values of the best, second

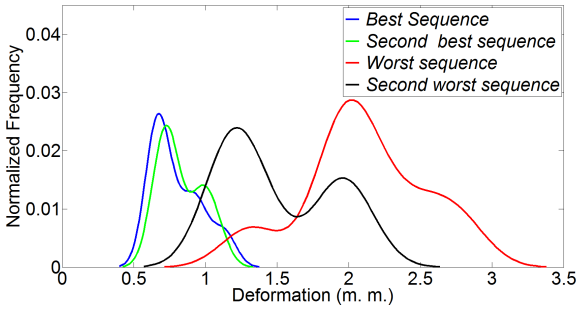


Fig. 17. Deformation.

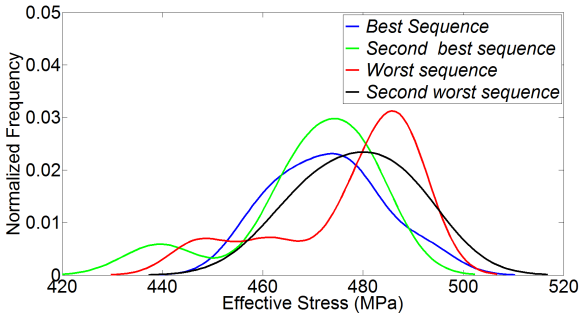


Fig. 18. Effective Stress.

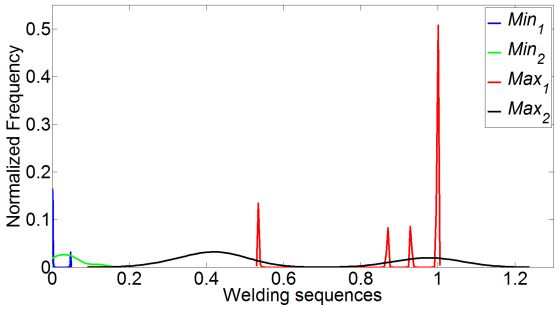


Fig. 19. Multi-Objective.

best, worst and second worst sequences respectively. Fig. 20, 21 and 22 demonstrate the deformation, effective residual stress and multi-objective values of these four sequences respectively in terms of the percentage if we consider the value of the worst sequence (red color bar) as 100%. Fig. 20, 21, and 22 illustrate that best sequence (blue color bar) achieves ∼80% maximum structural deformation, ∼15% maximum effective stress and

∼60% maximum multi-objective values over worst sequence (red color bar) respectively. Fig. 20, 21, and 22 also demonstrate that both best and second best sequences obtains substantial reduction of maximum structural deformation, effective stress and multi-objective values over worst and second worst sequences (red and black bars are much taller than blue and green bars). This result clearly demonstrates that welding sequence has significant effect on welding optimization technique.

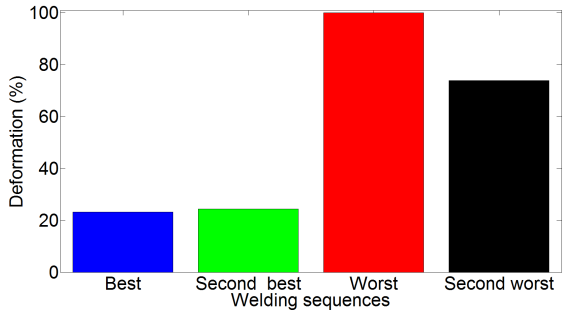


Fig. 20. Deformation.

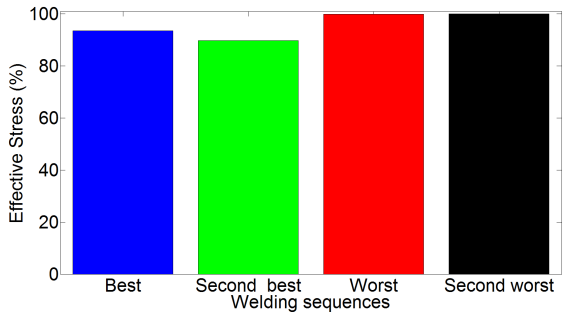


Fig. 21. Effective Stress.

6 Conclusion and Future Work

Structural deformation and effective residual stress defines a measurement of quality in terms of welded structures. In this research, we developed and implemented a multi-objective GA based on welding sequence optimization. Both structural deformation and effective residual stress have been combined. We exploited a multi-objective fitness

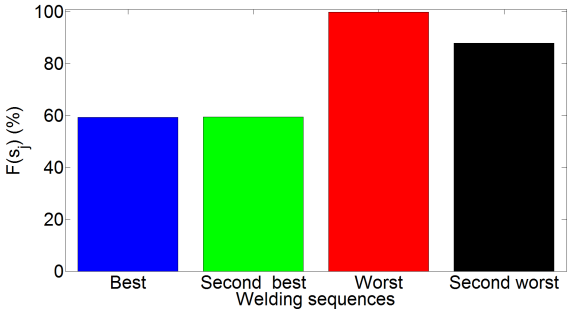


Fig. 22. Multi-Objective.

function which consists of the linear combination of the inverse of the structural deformation and effective residual stress have been linearly combined and the optimal weight between them are dynamically selected in an iterative manner ensuring an equitable treatment of both structural deformation and effective stress.

A thermo-mechanical FEA was used to compute the structural deformation and effective residual stress. The three sequential steps of GA: selection, crossover and mutation were tailored for welding sequence optimization to facilitate convergence and avoiding the repetition of the weld seam in the sequence.

An elitism selection approach was implemented by copying the three best individuals into the next generation to expedite the convergence as well as preserving the good chromosomes which have high probability to offer optimal solution. We implemented a sequential string searching algorithm to adjust the single point crossover algorithm for the welding sequence optimization to avoid the repetition of single bead into the welding sequence. For the similar reason, we only changed the direction of the welding seam instead the welding seam itself. We computed and executed minimum number of iterations necessary for finding the optimal solution of the GA based on the general Markov chain model of GA.

We carried out a simulation experiment on a mounting bracket which its design is widely used in vehicles and other applications. Experiments were conducted on a structure with eight weld seams. Results of Pareto front demonstrate that both structural deformation and effective residual stress cannot be simultaneously reduced. Experimental results illustrate that best welding sequence can reduce significant amount of structural deformation ($\sim$80%) and residual stress ($\sim$15%) over worst sequence.

This research launched multiple new directions to the welding sequence optimization. In future, we would like to incorporate other important characteristics to measure the quality of welded structures such as temperature, robot time and robot path for welding sequence optimization. Information about the structural deformation and effective residual stress after welding each seam in a sequence needs to be investigated.

References

1. Alexandrescu, A., Agavriloaei, I.: Determining the best mutation probabilities of a genetic algorithm for mapping tasks. The Bulletin of The Polytechnic Institute from Iaşi LVII(LXI), 10 (2011)
2. Aytug, H., J., K.G.: Stopping criteria for finite length genetic algorithms. INFORMS Journal on Computing 8(2), 183–191 (1996)
3. Aytug, H., J., K.G.: New stopping criterion for genetic algorithms. European Journal of Operational Research 126(3), 662–674 (2000), http://www.sciencedirect.com/science/article/pii/S0377221799003215
4. Chapple, A., Tahir, Z., Jardine, F.: Weld Distortion Optimisation using HyperStudy. The 8th UK Altair Technology Conference 2013 pp. 1–13 (2013)
5. Derlukiewicz, D., Przybyek, G.: Chosen aspects of {FEM} strength analysis of telescopic jib mounted on mobile platform. Automation in Construction 17(3), 278–283 (2008), http://www.sciencedirect.com/science/article/pii/S0926580507000714, 8th International Conference on Computer-Aided Engineering
6. Fu, G., Loureno, M.I., Duan, M., Estefen, S.F.: Influence of the welding sequence on residual stress and distortion of fillet welded structures. Marine Structures 46, 30–55 (2016), http://www.sciencedirect.com/science/article/pii/S0951833915001021
7. Gabli, M., Jaara, E.M., Mermri, E.B.: A genetic algorithm approach for an equitable treatment of objective functions in multi-objective optimization problems. International Journal of Computer Science 41(2) (2014)
8. Goldak, J., Asadi, M.: Challenges in verification of cwm software to compute residual stress and distortion in welds. In: ASME 2010 Pressure Vessels and Piping Conference: Volume 3. pp. 1–9 (2010)
9. Goldak, J.A., Akhlaghi, M.: Computational welding mechanics. Springer (2005)
10. Goldberg, D.E.: Genetic Algorithms in Search, Optimization and Machine Learning. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1st edn. (1989)
11. Hasançebi, O., Erbatur, F.: Evaluation of crossover techniques in genetic algorithm based optimum structural design. Computers & Structures 78(13), 435–448 (2000), http://www.sciencedirect.com/science/article/pii/S0045794900000894
12. Hussain, M.: Study of the dynamic behavior of a car body for mounting the rear axle. In: Proceedings of the European Automotive Congress (EAEC-ESFA). vol. 25 (2016)
13. Islam, M., Buijk, a., Rais-Rohani, M., Motoyama, K.: Simulation-based numerical optimization of arc welding process for reduced distortion in welded structures. Finite Elements in Analysis and Design 84, 54–64 (jul 2014), http://linkinghub.elsevier.com/retrieve/pii/S0168874X14000201
14. Jackson, K., Darlington, R.: Advanced engineering methods for assessing welding distortion in aero-engine assemblies. IOP Conference Series: Materials Science and Engineering 26, 012018 (2011)
15. Kadivar, M.H., Jafarpur, K., Baradaran, G.H.: Optimizing welding sequence with genetic algorithm. Computational Mechanics 26(6), 514–519 (2000)
16. Kaya, Y., Uyar, M., Tekin, R.: A novel crossover operator for genetic algorithms: Ring crossover. CoRR abs/1105.0355 (2011), http://dblp.uni-trier.de/db/journals/corr/corr1105.html#abs-1105-0355
17. Kim, K.Y., Kim, D.W., Nnaji, B.O.: Robot arc welding task sequencing using genetic algorithms. IIE Transactions 34(10), 865–880 (2002)
18. Koenig, A.: A study of mutation methods for evolutionary algorithms. CS 447 - Advanced Topics in Artificial Intelligence pp. 1–8 (2002), http://web.mst.edu/{~}tauritzd/courses/ec/fs2002/project/Koenig.pdf

19. Kumar, D.A., Biswas, P., Mandal, N.R., Mahapatra, M.M.: A study on the effect of welding sequence in fabrication of large stiffened plate panels. Journal of Marine Science and Application 10(4), 429–436 (2011)
20. Liao, Y.G.: Optimal design of weld pattern in sheet metal assembly based on a genetic algorithm. International Journal of Advanced Manufacturing Technology 26(5-6), 512–516 (2005)
21. Lindgren, L.E.: Computational welding mechanics. In: Lindgren, L.E. (ed.) Computational Welding Mechanics, pp. 31–46. Woodhead Publishing Series in Welding and Other Joining Technologies, Woodhead Publishing (2007), http://www.sciencedirect.com/science/article/pii/B9781845692216500045
22. Masubuchi, K.: Analysis of Welded Structures, vol. 3. Pergamon Press Ltd. (1980)
23. Mohammed, M.B., Sun, W., Hyde, T.H.: Welding sequence optimization of plasma arc for welded thin structures. In: WIT Transactions on The Built Environment. vol. 125, pp. 231–242 (2012), http://library.witpress.com/viewpaper.asp?pcode=OP12-020-1
24. Soni, N., Kumar, T.: Study of various mutation operators in genetic algorithms. International Journal of Computer Science and Information Technologies 5, 4519–4521 (2014)
25. Studniarski, M.: Stopping criteria for genetic algorithms with application to multiobjective optimization. In: Schaefer, R., Cotta, C., Kołodziej, J., Rudolph, G. (eds.) Parallel Problem Solving from Nature, PPSN XI: 11th International Conference. pp. 697–706. Springer (2010)
26. Subbiah, S., Singh, O., Mohan, S.K., Jeyaraj, A.P.: Effect of muffler mounting bracket designs on durability. Engineering Failure Analysis 18(3), 1094–1107 (2011)
27. Xie, L.S., Hsieh, C.: Clamping and welding sequence optimization for minimizing cycle time and assembly deformation. International Journal of Materials & Product Technology 17(5/6), 389–399 (2002)

Impreso en los Talleres Gráficos
de la Dirección de Publicaciones
del Instituto Politécnico Nacional
Tresguerras 27, Centro Histórico, México, D.F.
noviembre de 2017
Printing 500 / Edición 500 ejemplares