IAM'2019

Proceedings

2ND CONFERENCE ON INFORMATICS AND APPLIED MATHEMATICS

12-13 JUNE 2019

GUELMA UNIVERSITY

SPEAKER

PROF. MOHAMED BENMOHAMMED

IS A PROFESSOR IN COMPUTER SCIENCE AT UNIVERSITY OF CONSTANTINE 2, ALGERIA. HE RECEIVED THE PHD DEGREE FROM UNIVERSITY OF SIDI BEL-ABBES, ALGERIA, IN 1997. HIS RESEARCH INTERESTS INCLUDE: MICROPROCESSOR, COMPUTER ARCHITECTURE, EMBEDED SYSTEM AND COMPUTER NETWORKS.

EMBEDED SYSTEMS, INTERNET OF THINGS IOT VS SMART CITIES

ABSTRACT

THE SMART CITIES CONCEPT INTEGRATES INFORMATION AND COMMUNICATION TECHNOLOGIES (ICT) AND VARIOUS PHYSICAL DEVICES CONNECTED TO THE NETWORK (INTERNET OF THINGS IOT) TO OPTIMIZE THE EFFICIENCY OF URBAN OPERATIONS AND SERVICES AND CONNECT TO CITIZENS. THE SMART CITIES MARKET IS IN CONSTANTLY EVOLUTION. THIS MARKET TO REACH $1,400 BILLION BY 2020, SAYS A RECENT REPORT. BILL CLINTON WAS THE FIRST TO TALK ABOUT THE CONCEPT OF "SMART CITIES" IN 2005, EXPRESSING THAT CITIES ARE IN FACT ALREADY SMART BUT MUST NOW BECOME SUSTAINABLE. EMBEDED SYSTEMS (EMBEDDED CAMERAS, EMBEDDED SENSORS, ROUTERS, COMMUNICATION NETWORKS, ETC.) ARE A SPECIFIC TECHNOLOGY THAT HELPS CREATE SMART CITIES. THEIR AIM IS TO CREATE A DISTRIBUTED NETWORK OF INTELLIGENT SENSOR CORES THAT CAN MEASURE SEVERAL INTERESTING PARAMETERS FOR BETTER MANAGEMENT OF THE CITY. ALL DATA ARE TRANSMITTED IN REAL TIME TO THE CITIZENS OR AUTHORITIES CONCERNED.

Honorary Chairmen

- **Pr. Salah Ellagoune, rector of 8 may 1945 guelma university**
- **Pr. Athmane meddour, dean of MI/SM faculty**

Organisation Committee

- **Dr. Mohamed Nadjib KOUAHLA, General Chair**
- **Dr. Brahim Farou,**
- **Dr. Khaled halimi,**
- **Mr. Samir hallaci,**
- **Dr. zineddine kouahla**

Program Committee

Chairmen

Pr. Hamid Seridi	**University of Guelma, Algeria**
Pr. Muhammet Kurulay	**University of Yildiz Technical, Turkey**
Pr. M-Z Aissaoui	**University of Guelma, Algeria**

Members

Abdelmalek	Amine	University of Saida, Algeria
Abdelkrim	Amirat	University of Souk Ahras, Algeria
Ahmad Kamran	Malik	University of COMSATS, Pakistan
Ahmed	Azough	Sidi Mohamed Ben Abdellah, Morocco
Ahmed	Rebai	Esprit Engineering School of Sfax, Tunisia
Antonio	Guerrieri	National Research Council of Italy, Italy
Azzedine	Bilami	University of Batna, Algeria
Bachir	Boucheham	University of Skikda, Algeria
Baghdad	Atmani	University of Oran, Algeria
Basit	Raza	University of COMSATS, Pakistan
Benmohammed	Mohammed	University of Constantine, Algeria
Boufaida	Zizette	University of Constantine, Algeria
Bouhadada	Tahar	University of Annaba, Algeria
Brahim	Farou	University of Guelma, Algeria
Chemesse Ennehar	Bencheriet	University of Guelma, Algeria
Cherif	Foudil	University of Biskra, Algeria
Cherki	Daoui	University of Sultan Moulay Slimane, Morocco
Claudio	Savaglio	University of Calabria, Italy
Cyril	de Runz	University of Reims Champagne-Ardenne, France
Djakhdjakha	Lynda	University of Guelma, Algeria
Douadi	Bourouaieh	University of Guelma, Algeria

Elif Segah	Oztas	University of Karamanoglu Mehmetbey, Turkey
Elkamel	Merah	University of Khenchela, Algeria
Fateh	Ellaggoune	University of Guelma, Algeria
Fatih	Karaaslan	University of Yildiz Technical, Turkey
Giancarlo	Fortino	University of Calabria, Italy
Hamid	Benseridi	University of Setif, Algeria
Hamid	Seridi	University of Guelma, Algeria
Hamid	El Maroufy	University of Sultan Moulay Slimane, Morocco
Hasna	Bouazza	University of Oran, Algeria
Inaya	Lahoud	University of Galatasaray, Turkey
Ismahane	Souici	University of Jijel, Algeria
Kamal E	Melkemi	University of Biskra, Algeria
Karim	Bouamrane	University of Oran, Algeria
Maamar	Sedrati	University of Batna, Algeria
Mahmoud	Boufaida	University of Constantine, Algeria
Mebarka	Yahlali	University of Saida, Algeria
Mohamed	Nemissi	University of Guelma, Algeria
Mohamed Amine	Boudia	University of Saida, Algeria
Mohamed Amine	Ferrag	University of Guelma, Algeria
Mohamed Chaouki	Babahenini	University of Biskra, Algeria
Mohamed Chawki	Batouche	University of King Saud, Saudi Arabien
Mohamed Nadjib	Kouahla	University of Guelma, Algeria
Mohamed Reda	Hamou	University of Saida, Algeria
Mohamed-Khireddine	Kholladi	University of Eloued, Algeria
Nacereddine	Zarour	University of Constantine, Algeria
Nacira	Ghoualmi-Zine	University of Annaba, Algeria
Nadir	Farah	University of Annaba, Algeria
Nadjia	Benblidia	University of Blida, Algeria
Okba	Kazar	University of Biskra, Algeria
Said	Talhi	University of Batna, Algeria
Salim	Chikhi	University of Constantine, Algeria
Smaine	Mazouzi	University of Skikda, Algeria
Yacine	Lafifi	University of Guelma, Algeria
Zahia	Guessoum	University of Paris 8, France
Zineddine	Kouahla	University of Guelma, Algeria

Author Index

Keyword Index

Table of Contents

K-means and Sammon Mapping Visual Analytics Based Clustering Methods Applied to Corrupted Documents for Authorship Attribution

Zohra Hamadache
Faculty of Electronics and Computer Science
USTHB University
Algiers, Algeria
zohra.hamadache@yahoo.fr

Halim Sayoud
Faculty of Electronics and Computer Science
USTHB University
Algiers, Algeria
halim.sayoud@uni.de

***Abstract*—Visual Analytics (VA) can be described as the application of interactive techniques in order to communicate data. It combines reasoning analytics and interactive information visualization with human cognition and perception, visual intelligence, data mining and data management. To perform an efficient investigation and implement a consistent evaluation of visual analytics methodology, a suitable infrastructure in terms of database and software tool, is needed. In this study, we focus on the VA based Authorship Attribution (AA), applied on noisy text data. For that purpose, we combined different stylistic features with two VA based clustering algorithms. Hence, we used a dataset that contains several text documents written by 5 American Philosophers, with an average length of 850 words per text, which are scanned and then corrupted with different noise levels. Compared to Sammon Mapping technique, our tests show that when noise level goes from 0% to 7%, the use of the K-means with Character Bigrams provides a consistent clustering. Actually, this combination succeeded to detect the whole number of text documents by indicating a Clustering Recognition Accuracy (CRR) of 100% at all noise levels.**

Keywords— Visual Analytics, Authorship Attribution, Noisy Text, Clustering, K-means, Sammon Mapping, Artificial Intelligence.

I. INTRODUCTION

The identification of the author of a given document is the Authorship Attribution or stylometry [1, 2, 3]. It is a case of the application of stylistic properties of a written text. Thus it involves the use of clustering techniques with the aim of discriminating authors. For that reason, we combined AA with VA which is the combination of the reasoning analytics with interactive information visualization [4 , 5 , 6 , 7], in order to deal with a large amount of noisy, complex, and voluminous data. Accordingly, the originality of this investigation is the test of the performances of K-means and Sammon Mapping VA based techniques under noise when different authors are disputed.

The remainder of this paper is structured as follows. In Section 2, we describe the related work of AA using VA. The implemented methodology is detailed in Section 3. Section 4 offers the evaluation steps. In Section 5, we present the used corpus based on noisy text data. In Section 6, we present our experimental and statistical results. In Section 7, we offer the main important conclusions and suggest some future perspectives.

II. RELATED WORK

Centuries ago [1, 2], AA or author identification was an issue that concerned many researchers because of its important role in authentication. In this investigation, we proposed the association of VA with AA. For that purpose, we did a research to explore about this combination and we found only three works.

The first exploration was carried out by [8] with the intention of scrutinizing the performance of feature values across the text. They made a combination of an efficient VA technique with automatic literature analysis algorithms by computing a succession of feature values on different hierarchy levels for each text and presenting them to the user as characteristic fingerprint. They examined some literary features for the AA using a corpus which consists of a total number of 16 texts freely accessible from Project Gutenberg where 6 texts have been written by Jack London and 10 texts have been written by Mark Twain. Their novel VA technique indicated successful applications to known literature issues.

The second exploration was made by [9] using VA in the Christian religion. They described the use of the multivariate statistical technique in the stylometric study to investigate the AA of all the New Testament texts. They used the Johannine corpus where they took the 75 most frequent words for every column, and the original Greek where they took the 500 word samples. They found that there is a slight difference in the writing style of authors in all texts of the New Testament.

The third study was performed by [10] using VA in the ancient documents of the Islamic religion. They explored the AA of the Holy Quran - with 14 text segments - which corresponds to the holy words of ALLAH (God) and the Hadith - with 11 text segments - which corresponds to the statements said by the prophet Muhammad. In order to discriminate between both religious books, the VA system was based on the Fuzzy C-mean and the hierarchical clustering techniques

combined with 7 features. It should be noticed that before the classification procedure, features were normalized by Principal Component Analysis (PCA) reduction. The experimentation results indicated that the holy Quran and Hadith were written by two different authors.

A. *The Difference Between the Proposed Study and Previous Studies*

So far, no work using VA, on noisy data, exists for the best of our knowledge. That is, our work could be considered as a first investigation in AA using K-means, and Sammon Mapping under noisy environments.

III. Methodology

In order to identify authors, we implemented a methodology that consists of three fundamental steps: 1) Digitalization and extraction chain of noisy text documents using Optical Character Recognition (OCR); 2) features extraction and; 3) applying clustering algorithms.

A. *Digitalization and Extraction Chain of Noisy Text Documents Using OCR*

The transformation of a text document into an image needs to find innovative ideas regarding the categorization and accessibility of texts in the digital file. The digitalization and extraction chain of noisy text documents starts from how to convert the written text into an image, how to add salt and pepper noise to the scanned text, and ends with how to extract the text document from the corrupted digital file using an OCR tool [11]. The resulting corrupted text is then used as input file in the authorship attribution process.

B. *Features Extraction*

It consists of seeking the fingerprint of the authors. Feature extraction [12] plays an important role when discriminating authors because of its great impact in terms of reflecting the writing style to a great extent. After the preprocessing step, it is very important to extract the best features that perform ideally with the clustering technique to reach the highest accuracy level in clustering doubtful texts. In this context, the writing style of the author provides a very powerful support to characterize a discernible feature.

In this work, we tested the performance of the following text features: *Characters* [13]; *Character N-grams* [14]; *Words* [15]; and *Ending bigrams*.

C. *Applying Clustering Algorithms*

For AA, we used K-means and Sammon Mapping clustering approaches.

IV. Clusterin Approaches

A. *K-Means clustering*

This algorithm is based on the definition of member objects and a centroid or center for every cluster in the partition, which represents the median or mean point of a group of points to which the sum of distances from all objects in that cluster is minimized. These centroids must be positioned as much as possible distant from each other. Then, every point belonging to a given data must be taken and should be related to the nearest centroid [16, 17]. The K-means clustering results will change according to the selection of the distance metric. Therefore, it is very important to select the appropriate metric in order to get contrasted results.

In this study, we used the Cosine distance [18] as input parameter-which is one minus the cosine of the included angle between points - for K-means clustering technique, knowing that “a” and “b” are the two objects. The Cosine distance is defined as follows:

$$Dist\,(a,b) = 1 - \frac{\sum_i \left(a_i\ b_i\right)}{\sqrt{\left(\sum_i a_i^2\right)\left(\sum_i b_i^2\right)}} \quad (1)$$

B. *Sammon Mapping clustering*

This technique is an unsupervised nonlinear model proposed by John W. Sammon in 1969 [19]. It is based on the transformation of data from a high-dimensional space to a space with lower dimensionality in order to visualize high-dimensional data. The objective of Sammon's mapping clustering technique is the minimization of the Sammon's error or Sammon's stress function described by the following expression:

$$E = \frac{1}{\sum_{i<j} d_{ij}^*} \sum_{i<j} \frac{\left(d_{ij}^* - d_{ij}\right)^2}{d_{ij}^*} \quad (2)$$

Where d_{ij}^* represents the distance between i^{th} and j^{th} objects in the original space, d_{ij} represents the distance between their projections.

V. Evaluation

Clustering methods evaluation was performed using the visual assessment by observing visually only the clusters organization, and the statistical computation of the Clustering Recognition Rate (CRR) values computed as follows.

$$\text{Clustering Error Rate (\%)} = \left(\frac{\text{Number of samples badly clustered}}{\text{Total number of text segments}}\right) * 100 \quad (3)$$

$$\text{Clustering Recognition Rate (\%)} = 100\% - \text{Clustering Error Rate} \quad (4)$$

VI. OCR5P Corpus

Since it was difficult to find a corpus containing potential authors with disturbed text documents written in the same period of time, with the same language and topic, we decided to construct our own dataset, which we called "Optical Character Recognition of 5 Philosophers" (i.e. OCR5P), with the aim of identifying authors under noisy conditions. The collection of this database was carried out in 2016 and comprises 5 American Philosophers, where each author possesses 5 text segments with an average size of approximately 850 words per text.

During our experiments of AA, we made several tests to see whether the text segments were correctly attributed or not after clustering process. Each text document was processed and disturbed into different noise levels as follows: 1) the text segments of all authors were scanned. 2) After that, a granular noise called "salt and pepper noise" was added to the obtained image data by increasing the noise levels from 0% to 7% as illustrated in Table 1.

TABLE I. Summary of the OCR5P Corpus

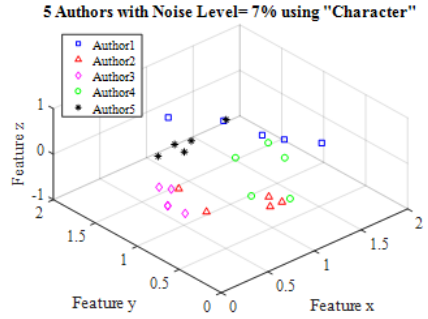
Authors	**Number Of Words Per Text (Noise Level From 0% To 7%)**				
	Text 1	***Text 2***	***Text 3***	***Text 4***	***Text 5***
Author1: Chauncey Wright	723	769	779	793	813
Author2: Corliss Lamon	824	824	838	801	829
Author3: Henri Bergson	981	965	972	957	933
Author4: Michael James	923	949	858	888	839
Author5: Solomon Gabriol	878	855	855	871	873

VII. Experimental Results and Discussions

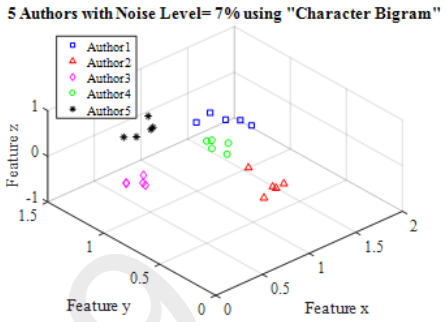
In this section, we are going to evaluate the developed VA based clustering methods with the purpose of recognizing the authors of noisy text documents. Graphical visualizations are gotten using the following text features: 1- Character, 2- Character Bigram, 3- Character Trigram, 4- Word and 5- Ending Bigram.

A. *K-means Results*

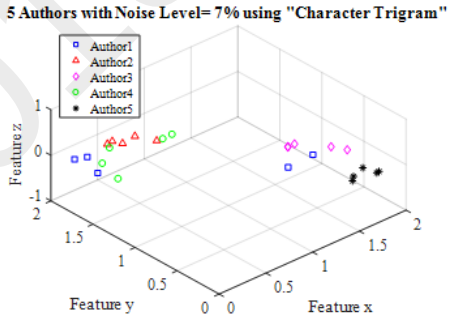
Results at 7% of noise are represented in Fig. 1. The input parameter is: 1- *cosine* distance metric.



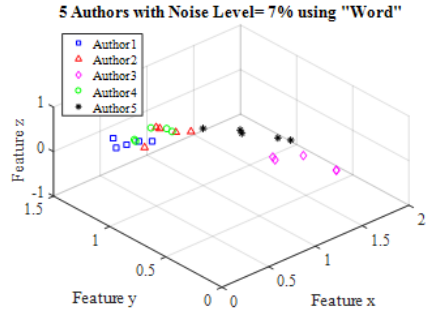
(a)



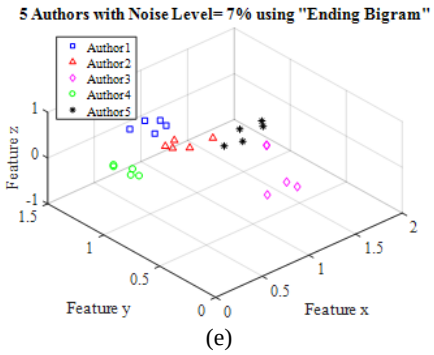
(b)



(c)



(d)



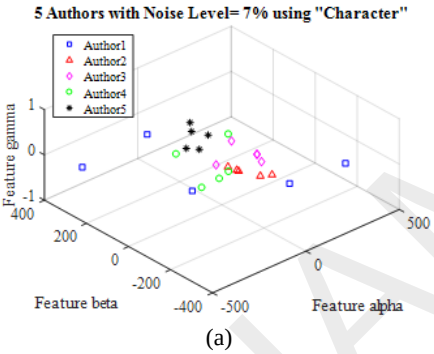
(e)

Fig. 1. Clustering Results obtained by the K-means technique after PCA reduction for 5 Authors with 5 text segments using as features: (a) Character, (b) Character Bigram, (c) Character Trigram using the 5000 most common features, (d) Word and (e) Ending Bigram. Noise Level=7%.

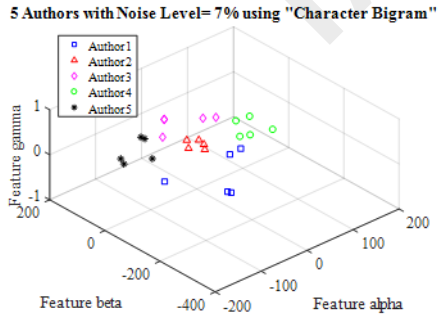
The K-means performance is evaluated when no intersection is seen between all text documents with regards to their similar properties. Accordingly, interesting results are provided by "*Character Bigrams*" at a noise level of 7%. In fact, all similar text segments having the same color and symbol are grouped together in a consistent way. This involves the identification of the five different styles.

B. Sammon Mapping Results

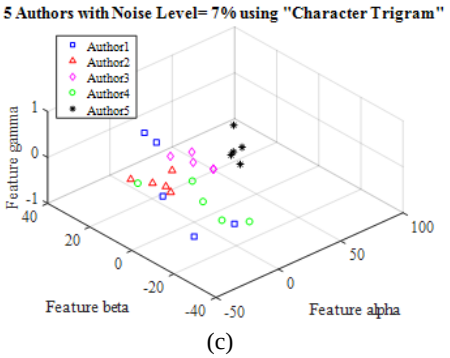
Results at 7% of noise are represented in Fig. 2.



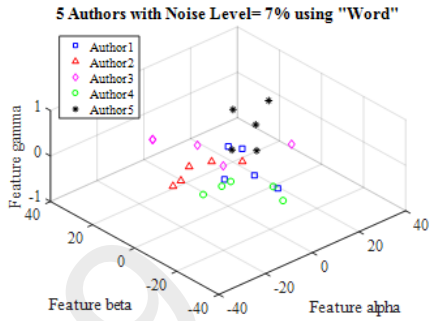
(a)



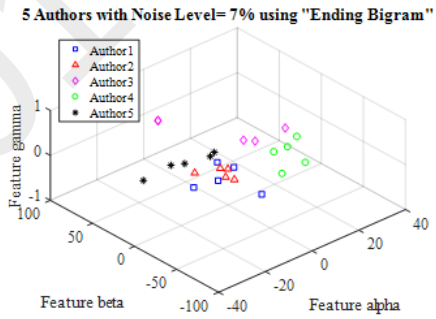
(b)



(c)



(d)



(e)

Fig. 2. Clustering Results obtained by the Sammon Mapping technique after PCA reduction for 5 Authors with 5 text segments using as features: (a) Character, (b) Character Bigram, (c) Character Trigram using the 5000 most common features, (d) Word and (e) Ending Bigram. Noise Level=7%.

For Sammon Mapping function case, the different writing styles are distinguished according to the three kept features represented by alpha, beta and gamma. By following that logic, three styles are detected using: "*Character Bigrams*" at a noise level of 7%.

C. CRR Results

To see the variation of the CRR *vs.* noise, we measured the CRR values showed in Table 2 for the K-means clustering and in Table 3 for the Sammon Mapping clustering. These computations were carried out up to the noise level of 7% and for different text features.

TABLE II. THE CRR UP TO 7% OF NOISE FOR DIFFERENT FEATURES USING K-MEANS

Features	Noise Level (%)							
	0	***1***	***2***	***3***	***4***	***5***	***6***	***7***
Character	88	88	96	88	92	80	88	68
Character Bigram	100	100	100	100	100	100	100	100
Character Trigram	76	88	88	96	84	76	84	76
Word	92	96	100	100	92	88	92	76
Ending Bigram	96	96	96	100	100	100	92	92

TABLE III. THE CRR UP TO 7% OF NOISE FOR DIFFERENT FEATURES USING SAMMON MAPPING

Features	Noise Level (%)							
	0	***1***	***2***	***3***	***4***	***5***	***6***	***7***
Character	84	76	80	96	88	76	76	60
Character Bigram	92	96	92	84	88	92	76	88
Character Trigram	96	88	88	96	88	88	76	80
Word	76	96	80	84	88	72	68	68
Ending Bigram	88	80	88	84	80	80	76	72

According to the statistical results, we can easily notice that at all noise levels, the combination of K-means with "*Character Bigrams*" yields interesting authorship attribution. In fact, it indicates that the CRR reached an accuracy of **100 %** which involves a consistent identification of the author under noise.

VIII. CONCLUSION AND FUTURE WORK

In this survey, we tested the robustness of K-means and Sammon mapping clustering techniques on a corpus with different levels of noise. This was performed by integrating different stylistic features. We showed that the author of a written text can be correctly recognized under noise using K-means clustering since it succeeded to recognize all text documents at all noise levels. Therefore, this clustering algorithm seems to be useful for authorship attribution with visual analytics under noisy conditions. We recall that the used corpus contains 25 text documents written in English by 5 American philosophers that are scanned to get images, and then corrupted with different noise levels by using the "salt and pepper" noise. Finally, they are transformed into text format using an OCR system. After several experiments, we noticed that the combination of the VA based K-means clustering technique with "Character Bigrams" performed better than Sammon Mapping, since very interesting authorship attribution results are obtained. Actually, an accuracy of 100% is reached when noise goes from 0% to 7%. Besides, the use of Sammon Mapping with all features showed that the recognition accuracy decreases when the noise level increases. This involves that this technique fails to detect the author of a corrupted document which is maybe due to the alteration of successive characters that form the same word, while minimizing the quantity of details during the corruption process. Therefore, the most efficient combination of classifier/feature, which appeared robust and accurate in AA is: K-means/ Character Bigrams.

As future work, we suggest testing the scalability of the K-means technique, to see whether it can work well on other datasets or on much more data. Besides, since the recognition accuracy of K-means is 100% at 7% of noise, it will be very interesting to investigate about the limit. For example, what happens if the noise level is, 70%? On the other side, we suggest to improve the Sammon Mapping performance by offering new input parameters, or thresholds, or new text features to create new combinations of Sammon Mapping / feature, to enable the human analyst to interact more easily with more complex and noisy text data.

REFERENCES

[1] S. E. De Morgan, "Memoir of Augustus de Morgan," London: Longmans, Green, pages 474, 1882. https://archive.org/details/memoirofaugustus00demorich.

[2] T. C. Mendenhall, "The Characteristic Curves of Composition," Science, Vol. 9, no. 214, pp. 237-249, 1887. http://www.jstor.org/stable/pdf/1764604.pdf?_=1469971242881.

[3] M. Piasecki, T. Walkowiak, and M. Eder, "Open Stylometric System WebSty: Integrated Language Processing, Analysis and Visualisation," CMST 24(1), pp. 43–58, 2018. http://cmst.eu/wp-content/uploads/files/10.12921_cmst.2018.0000007_PIASECKI_c.pdf.

[4] M. Song, Z. Pang, and E. Haihong, "A New Visual Analysis Approach to the High Dimensional Data," IOP Conf. Series: Journal of Physics: Conf. Series 1098 (2018), pp. 1 - 7, 2018. http://iopscience.iop.org/article/10.1088/1742-6596/1098/1/012008/pdf.

[5] D. Chen, B.M. Sanz, and E. Zhao, "Visual Analytics in the Public Sector: An Analysis on Diversities and SimilarIties of London's Wards," International Conference on Big Data Analytics, Data Mining and Computational Intelligence 2018 (BigDaCI 2018), pp. 1 - 6, 2018. http://researchopen.lsbu.ac.uk/2275/.

[6] Z. Sahaf, H. Hamdi, R.C.R. Mota, M.C. Sousa, and F. Maurer, "A Visual Analytics Framework for Exploring Uncertainties in Reservoir Models," In Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2018), pp. 74-84, 2018. http://www.scitepress.org/Papers/2018/66085/66085.pdf.

[7] D. Sacha, M. Kraus, J. Bernard, M. Behrisch, S. Schreck, Y. Asano, and D.A. Keim, "SOMFlow: Guided Exploratory Cluster Analysis with Self-Organizing Maps and Analytic Provenance," IEEE Transactions on Visualization & Computer Graphics, vol. 24, no. 1, pp. 120-130, 2018.

[8] D. A. Keim, and D. Oelke, "Literature Fingerprinting: A New Method for Visual Literary Analysis," IEEE Symposium on Visual Analytics Science and Technology (VAST 2007), pp. 115-122, 2007.

[9] H. Erwin, and M. Oakes, "Correspondence Analysis of the New Testament," Workshop on Language Resources and Evaluation for Religious Texts, pp. 1-8, 2012. http://pers-www.wlv.ac.uk/~in4326/papers/oakes_lrec_cam3.pdf.

[10] H. Sayoud, "A Visual Analytics based Investigation on the Authorship of the Holy Quran," In Proceedings of the 6th International Conference on Information Visualization Theory and Applications (IVAPP-2015), pp. 177-181, 2015. http://www.ivapp.visigrapp.org/?y=2015.

[11] L. Eikvil, "OCR - Optical Character Recognition," 1993 - academia.edu, pp. 1-35, 1993. https://www.nr.no/~eikvil/OCR.pdf.

[12] P. Juola, "Authorship Attribution," Foundation and Trends in Information Retrieval, Vol. 1, no. 3, pp. 233 - 334, 2006. http://www.mathcs.duq.edu/~juola/papers.d/fnt-aa.pdf.

[13] E. Stamatatos, "A Survey of Modern Authorship Attribution Methods," Journal of the American Society for Information Science and Technology, Vol 60, pp. 538- 556, 2009. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.207.3310&rep=rep1&type=pdf.

[14] D. Jurafsky, and J. H. Martin, "N-Grams," Speech and Language Processing, pp. 1 - 28, 2014. https://lagunita.stanford.edu/c4x/Engineering/CS-224N/asset/slp4.pdf.

[15] E. Stamatatos, N. Fakotakis, and G. Kokkinakis, "Automatic Authorship Attribution," Proceedings of EACL '99, pp. 158- 164, 1999.

[16] D. Agnihotri, K. Verma, and P. Tripathi, "Pattern and Cluster Mining on Text Data," 2014 Fourth International Conference on Communication Systems and Network Technologies (CSNT), pp. 428 - 432, 2014.

[17] R. K. Mishra, K. Saini, and S. Bagri, "Text document clustering on the basis of inter passage approach by using K-means," International Conference on Computing, Communication & Automation (ICCCA), pp. 110 - 113, 2015.

[18] J. Mingzhe, and J. Minghu, "Text Clustering on Authorship Attribution Based on the Features of Punctuations Usage," ICSP2012 Proceedings, pp. 2175– 2178, 2012.

[19] Jr. J. W. Sammon, "A nonlinear mapping for data structure analysis," IEEE Transactions on Computers, pp. 401-409, 1969.

Simple and Efficient Image Encryption Scheme based on Recursive Property and Plain Image-Chaotic Map Dependency

Tewfik Bekkouche
Dept. of electronics
(Faculty of technology)
ETA Laboratory
(University of Bordj bouarreridj)
Bordj bouarreridj 34000, Algeria
bekkou66@hotmail.com

Nacira Diffellah
Dept. of electronics
(Faculty of technology)
ETA Laboratory
(University of Bordj bouarreridj)
Bordj bouarreridj 34000, Algeria
nacirapush@gmail.com

Salah Mokhnache
Dept. of technology
(Faculty of technology)
(University of Setif1)
Setif1 19000, Algeria
mokhnachesalah@yahoo.fr

***Abstract*—In this paper, we propose an efficient symmetric image encryption scheme by using permutation-diffusion architecture. In order to resist to chosen-plaintext and chosen-ciphertext attacks a dependency is introduced between plain image and chaotic map both in permutation and diffusion processes. It consists firstly of scrambling the image chaotically then, we perform the bitxor operation element by element between each pixel of the scrambled image and its corresponding pixel in the same position of a chaotically generated image before performing recursively another bitxor operation between two consecutive resulting pixels following a given pattern. Computer simulation and security analysis confirm the efficiency of our proposed image encryption scheme in terms of histogram analysis, data loss, adjacent correlation and sensitivity tests.**

Keywords—permutation-diffusion, chaotic map, recursive property, dependency

I. INTRODUCTION

With the increasing technological progress experienced in the field of communication of information in the various social networks and on the Internet, this information is not immune from corruption and unauthorized manipulation. Their protection becomes essential. The image is part of this information flow. To protect it, several image encryption algorithms have emerged which are based on the permutation -confusion architecture [1]. If the permutation process consists in changing the positions of the pixels, the diffusion one ensures the change of their values by using the xor operator. Despite the effort made in the application of these algorithms, the results obtained will remain far from those expected, the use of chaotic maps in the field of image encryption in 1989 by Matthews [2] has given considerable support by giving an enhanced results which is mainly due to the chaotic properties of those chaotic maps as the high sensitivity of their parameters (initial value and control parameter) and since several algorithms have been designed based on chaos[3-8]. Although the results obtained by introducing chaos seem attractive and interesting, these encryption systems still remain vulnerable to a few attacks [9-16].

In this context and to remedy the problem mentioned above, we propose in this paper a method of image encryption based essentially on two main ideas, the first is to calculate the pixels average value of the original image and inject it into the parameters of the chaotic map jointly in the phase of permutation and diffusion, those parameters which become modulated in the rhyme of the plain image, the second consists to generate chaotically another image and performing recursively bit xor operation element by element between each pixel of this image and its corresponding pixel in the same position of plain image.

In the numerical methods, the main serious drawback of any recursive approach is the accumulation and propagation of the error. In contradiction with these methods, encryption techniques strongly search for any approach having this drawback or property. Therefore, we beneficially exploit this property of a recursive approach in images encryption to achieve the desired dependency between the pixels and hence ensure the accumulation and propagation of the error to all pixels in the case of any wrongness in the decryption key.

To achieve this purpose our paper is organized as follows: in section 2, we detailed both the proposed encryption and decryption schemes, simulation results and security analysis are discussed in section 3 and some concluding remarks are given in section4.

II. PROPOSED ENCRYPTION/DECRYPTION METHOD

A. PLCM chaotic map

Chaotic maps are known to have attractive cryptographic properties such as high sensitivity to their initial parameters, ergodicity, and pseudo-randomness. In our proposed encryption scheme, we exploit the piecewise linear chaotic map (PLCM) proposed in [17] and expressed iteratively as:

$$z_{k+1} = F(z_k, \lambda) = \begin{cases} \frac{z_k}{\lambda}, & 0 \le z_k < \lambda \\ \frac{z_k - \lambda}{0.5 - \lambda}, & \lambda \le z_k < 0.5 \\ F(1 - z_k, \lambda), & 0.5 \le z_k < 1 \end{cases} \quad (1)$$

Where $z_0$the initial is condition parameter and $\lambda \in (0,0.5)$ is the control parameter.

B. Encryption scheme

The encryption process is detailed in the following steps:

1- Let P be the original image of size $(N \times N)$ .We calculate firstly the pixels average value expressed by:

$$M = \frac{\sum_{ij} \mathrm{P}}{N \times N \times 255} \quad (2)$$

2- Resize the input image into a vector i of length $1 \times (N \times N)$.

3- Generate a chaotic vector z, $\{z_k, k = 1, 2, 3, \dots, N \times N\}$ using the PLCM given by Eq. (1) with the parameters $\{z_0 + 0.1M, \lambda_0 + 0.1M\}$.

4- Sort the vector **z** into an ascending order to form a vector **y** and then form a permutation map vector **m** such as m_k is

the position of the element y_k in the vector **z**, i.e., $\{y_k = z_{m_k}, k = 1,2,3,\dots,N \times N\}$.

5- Scramble the vector **i** using the permutation map vector **m** to form a vector **s** such as s_k is the $(m_k)^{\text{th}}$ element of **i**, i.e., $\{s_k = i_{m_k}, k = 1,2,3,\dots,N \times N\}$.

6- Generate another chaotic vector $\boldsymbol{x}$, $\{x_k, k = 1,2,3,\dots,N \times N\}$ using the PLCM given by Eq. (1) with the parameters $\{z_1 + 0.1M, \lambda_1 + 0.1M\}$, then converted it to eight-bit integer.

7- Perform the bit wise-xor operation recursively to form a vector v_{cry_k} according to the following formula:

$$v_{cry_k} = \begin{cases} s_k \oplus x_k \oplus r & k = 1 \\ s_k \oplus x_k \oplus v_{cry_{k-1}} & k = 2,3,\dots,N \times N \end{cases} \quad (3)$$

Where $r = round(255 * z_{(N \times N)})$, i.e., the random integer r is chosen to be the last element of the chaotic vector **z** converted to an eight-bit integer.

8- Finally, the vector v_{cry_k} is reshaped to obtain encrypted image designed by img_{cry}.

C. Decryption scheme

The decryption process takes exactly the steps of encryption process in inverse manner to obtain decrypted image img_{decry}.

III. Simulation Results and Security Analysis

Results of simulations are performed under environment MATLAB version 7.10.0.499 (R2010a), the used test images are those of Lena and Barbara of size (256×256), Living room and Clown of size (512×512). The two PLCM chaotic sequences having the following parameters:

$$(z_0 = 0.1428; \lambda_0 = 0.2567), (z_1 = 0.2428; \lambda_1 = 0.2867).$$

To evaluate the proposed method (TABLE I), we have used different metrics: the peak signal-to-noise ratio (PSNR) - the mean square error (MSE) and the standard correlation coefficient which are widely defined in previous works [5].

Figure.1 illustrates the used test images of (Lena, Barbara, Living room and Clown) and their histograms.

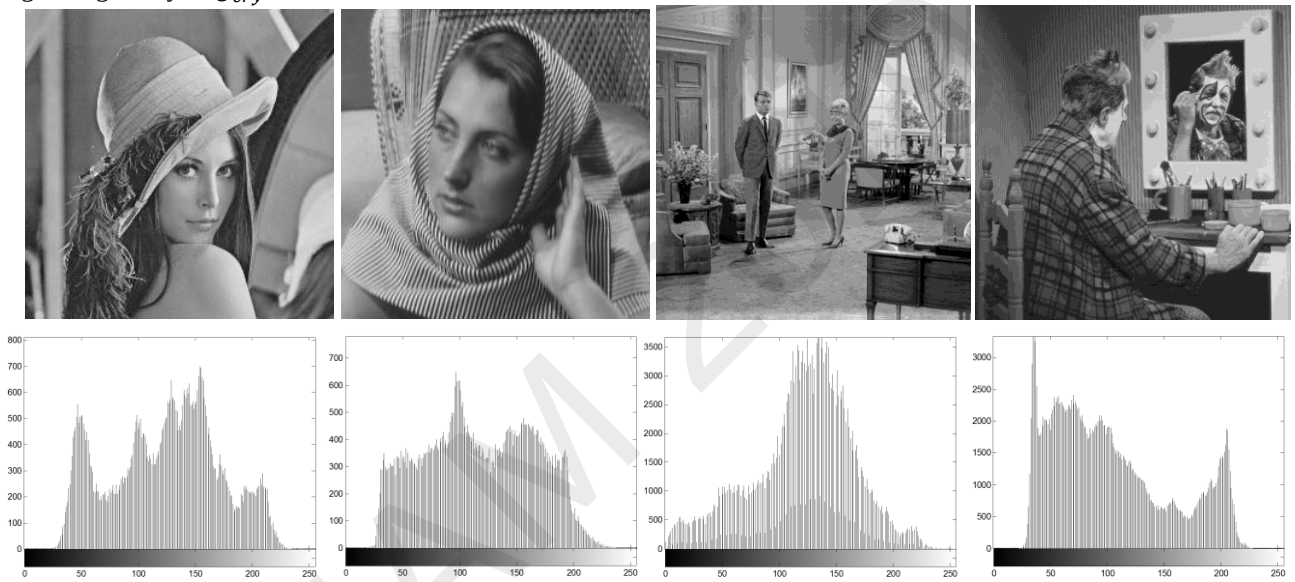


Fig. 1. Lena ,Barbara, Living room and Clown test images and their histograms.

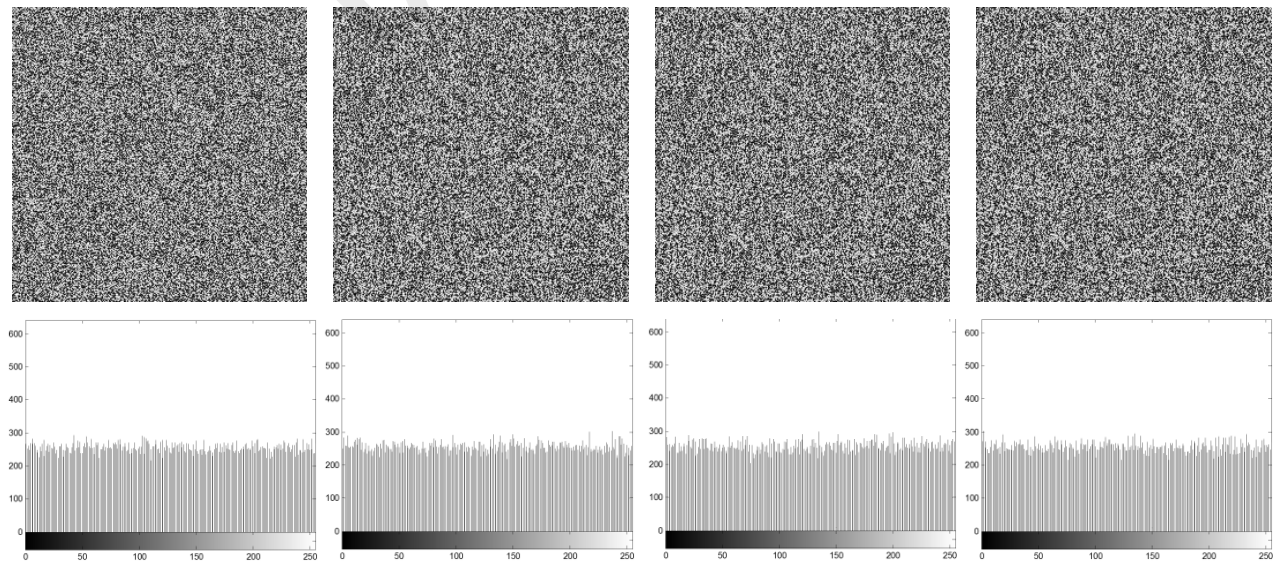


Fig. 2. Encrypted images of Lena , Barbara, Living room and Clown and their histograms.

Figure.2 illustrates encrypted Lena and Barbara and their histograms.

TABLE I. PSNR AND CORRELATION COEFFICIENT METRICS

Image	PSNR, dB			Correlation		
	Proposed	***[4]***	***[5]***	***Proposed***	***[4]***	***[5]***
Lena	9.2175	9.9013	9.2219	-0.0040	0.0071	-0.0059
Barbara	9.1359	9.1444	9.1835	-0.0053	0.0016	0.0059
Living room	9.3785	9.7103	9.3962	-0.0026	0.0020	0.0019
Clown	8.6713	9.3621	8.6766	0.0010	0.0040	0.0036

The results obtained for the PSNR and correlation coefficient are summarized in Table I for different methods and test images. It is clear from this table that the proposed method outperforms the methods presented in Refs. 4 and 5 in terms of PSNR, whereas results of Ref. 4 and Ref. 5 provide better with Barbara and Living room images in terms of the correlation coefficient.

A. Analysis of histogram

As shown in Fig. 2, starting from different histograms for different original images of Lena, Barbara, Living room and Clown, we find the same histograms of their encrypted images that look like a uniform white noise, this confirms that the proposed encryption method brings all original images back to encrypt images with the same histograms, which prevents attackers from deriving any information that may reveal the encryption operation and therefore, we can conclude that our approach resists attacks by analyzing histograms.

B. Entropy analysis

If histogram analysis simply shows the result qualitatively, entropy of information gives quantitative analysis as a measure of disorder that can quantify histogram uniformity. So for grayscale image encryption, the value of the entropy must be very close to 8, because if the entropy is less than 8, there are degrees of predictability, so we cannot provide security against statistical analysis. The entropy should ideally be 8 [4].

Entropy information is defined as:

$$H(I) = \sum_{i=0}^{F-1} p(i) log_2 \frac{1}{p(i)} \qquad (4)$$

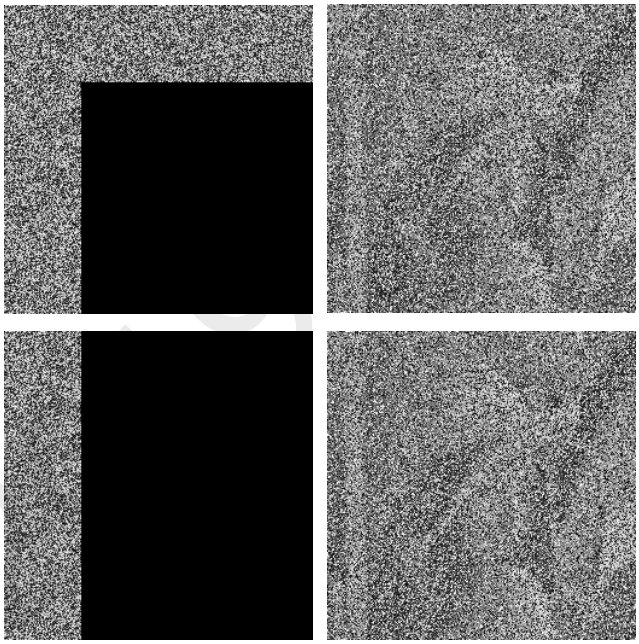
$P(i)$ represents the probability of pixels with the value equal to i (from o to $F-1$). According to table II bellow, information entropy has been increased with the encryption system and the information entropy of the encrypted image is almost near to 8. This confirms that the pixel values after encryption process seems random, which is sufficient secure for information leakage (TABLE II).

TABLE II. INFORMATION ENTROPY ANALYSIS

File name	Information entropy of original image	Information entropy of encrypted image
Lena	7.4429	7.9973
Barbara	7.4948	7.9968
Living room	7.2952	7.9993
Clown	7.3851	7.9994

C. Loss data

To test the resistance of the encryption method to loss data, we consider the case where a part of the pixels of the encrypted image was lost during transmission. From results of simulation shown in Fig .3, we note that due to the different percentages of data losses reaching the encrypted image up to a percentage of (75%), the information on the decrypted image is not totally eliminated; indeed the decrypted image remains visible and identifiable with the naked eye despite being noisy. Consequently, these results prove the robustness of the method toward data loss test and demonstrate its resistance to transmission errors.



(a) (b)

Fig. 3. Loss data test: (a) Encrypted Lena with different loss data (b) Corresponding decrypted Lena.

D. Correlation analysis

TABLE III. ADJACENT PIXELS CORRELATION ANALYSIS

Corrélation coefficient Original image/Encrypted image			
Direction	***Proposed method***	***[4]***	***[5]***
Horizontal	0.9258/0.0018	0.9258/-0.0105	0.9258/0.0027
Vertical	0.9593/-0.0022	0.9593/-0.0009	0.9593/0.0026
Diagonal (lower left to top right)	0.9258/-0.0008	0.9258/0.0029	0.9258/0.0004
Diagonal (lower right to top left)	0.9037/0.0007	0.9037/0.0079	0.9037/0.0024

To check the degree of destruction of dependency between adjacent pixels on an encrypted image by the proposed method, we randomly took a sample of 1000 adjacent pixels from this image; we measured the correlation coefficient between pixels in the three directions (vertical, horizontal, diagonal) and compare these measurements with those of the corresponding original image. (Table III) summarizes the correlation coefficient

measurements of the encrypted images and their original images in the three directions.

The correlation coefficients measured for the original images are close to 1, whereas the correlation coefficients of the encrypted images approach 0. It is deduced that the encryption has considerably attenuated the correlation between the pixels of the encrypted image.

Fig .4 shows respectively the correlation distributions of the horizontal, vertical and diagonal adjacent pixels of the original image and the encrypted image. This figure confirms the results in Table III, since the pixel intensity distribution of the original image focuses on the diagonal, so the pixels are highly correlated, while those in the encrypted image are uncorrelated and have a uniform distribution.

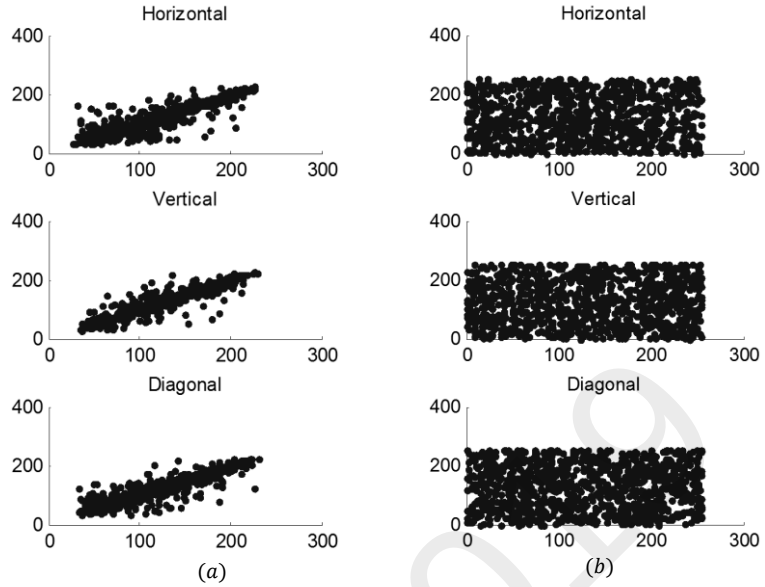


Fig. 4. Correlation distributions of the horizontal, vertical and diagonal adjacent pixels of (a) Lena image and (b) the encrypted Lena image.

E. Sensitivity analysis

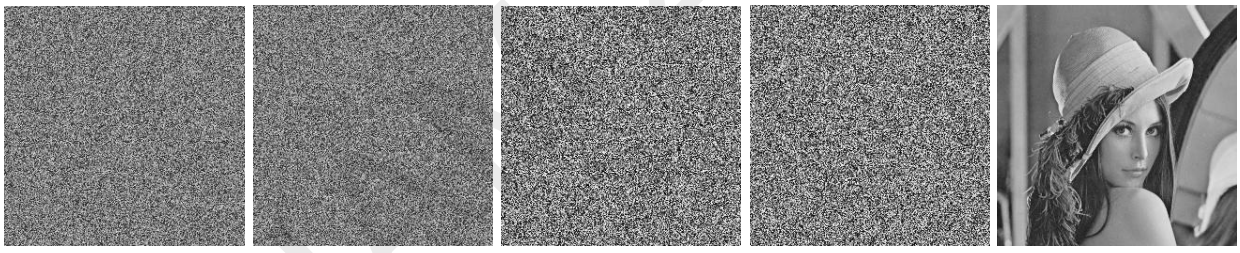


Fig. 5. Sensitivity test: Decrypted Lena with (a) $\lambda'_{0=}\lambda_0 + 0.1M + 10^{-16}$ (b) $z'_0 = z_0 + 0.1M + 10^{-16}$ (c) $\lambda'_{1=}\lambda_1 + 0.1M + 10^{-16}$ (d) $z'_1 = z_1 + 0.1M + 10^{-16}(e)$Correct key.

Let $k_1$be the encryption key of the proposed method, which is composed of the parameters of the two PLCM chaotic sequences $k_1(\lambda'_0, z'_0, \lambda'_1, z'_1)$, the corresponding decryption key is designed by $k_2(\lambda''_0, z''_0, \lambda''_1, z''_1)$. If the encryption key is exactly the decryption one i.e., $(k_1 = k_2)$ the decrypted image is identical to the original image Fig. $5(e)$.To test the sensitivity of the encryption key, we make a minor variation in one parameter of the order of 10^{-16} and we set the others parameters at their fixed values . The different cases are depicted in Fig. 5. $(a), (b), (c), (d)$. The results of simulations obtained show that the limit of the appearance of the image decrypted in the clear is of the order of 10^{-17}, this confirms the high sensitivity of the proposed method which is of the order of 10^{-16} for the different parameters of the encryption key.

F. Differential attack

To calculate the influence of a change of a single pixel in the original image and its repercussion on the corresponding encrypted image for any encryption algorithm, two quantities can be used, (NPCR)(number of pixels change rate), and (UACI)(Unified average changing intensity) defined by the following formulas:

$$NPCR = \frac{\sum_{i,j} D(i,j)}{M \times N} \times 100\% \quad (5)$$

$$UACI = \frac{1}{M \times N}\left[\sum_{i,j} \frac{C_1(i,j) - C_2(i,j)}{255}\right] \times 100\% \quad (6)$$

where $C_1(i,j)$ and $C_2(i,j)$ are the encrypted images corresponding to the original image C before and after modification of a single pixel in the original image, D being

a binary matrix having the same size $M \times N$ as the original image which is defined as follows:

$$D(i,j) = \begin{cases} 0 & C_1(i,j) = C_2(i,j) \\ 1 & C_1(i,j) \neq C_2(i,j) \end{cases} \quad (7)$$

The NPCR measures the percentage of the number of different pixels between the two images C_1 and C_1relative to the total number of pixels, while the UACI measures the average of difference intensity between two images C_1 and C_2.

TABLE IV. NPCR AND UACI PERFORMANCE

Image	With dependency		Without dependency	
	NPCR	***UACI***	***NPCR***	***UACI***
Lena	99.6292	33.4952	90.6143	45.4678
Barbara	99.6063	33.3577	90.6143	22.7136
Living room	99.6090	33.5041	36.2572	18.2166
Clown	99.6132	33.4265	36.2572	18.1935

According to TABLE IV, we can clearly observed the effect of dependency in improvements of NPCR and UACI performance

G. Key space

According to the previous results obtained in the test of sensitivity, we have reached a precision of 10^{+16} for each parameter of the encryption key, therefore, we conclude that the key space is evaluated at $10^{16\times4} = 2^{192}$, which is widely sufficient compared to the value required in cryptosystems[18].

IV. CONCLUSION

In this paper, we have shown that by introducing a dependency between plain image and PLCM chaotic map parameters, the performance of NPCR and UACI can significantly be improved. More ever, we have also shown that farther improvements in terms of sensitivity and robustness toward decryption attacks can be achieved by applying recursive property. Computer simulations confirm the simplicity and effectiveness of the proposed encryption method and outperform the existing chaos encryption systems.

REFERENCES

[1] C. E. Shannon, "Communication Theory of Secrecy Systems," Bell System Technical Journal, vol. 28 **(**1949), pp. 656–715.

[2] R. Matthews, "On the derivation of a chaotic encryption algorithm," Cryptologia, vol. 4 (1989), pp. 29–42.

[3] R. Parvaz and M. Zarebnia, "A combination chaotic system and application in color image encryption," Optics and Laser Technology, vol. 101 (2018), pp. 30–41.

[4] Lu Xu, Xu Gou, Zhi Li, Jian Li "A novel chaotic image encryption algorithm using block scrambling and dynamic index based diffusion, " Optics and Lasers in Engineering, vol. 91 (2017), pp. 41–52.

[5] A. Beloucif, O. Noui, L. Noui, "Design of a tweakle image encryption algorithm using chaos-based scheme," International journal of information and security, vol. 8 (2016), pp. 205–220.

[6] Chunyan.Han, "An image encryption algorithm based on modified logistic chaotic map," optik, vol. 181 (2019), pp. 779–785.

[7] H.Zhongyun, J.Fan, X.Binxuan, H.Hejiao, "2D Logistic-Sine-coupling map for image encryption," Signal processing, vol. 149 (2018), pp. 148–161.

[8] B. Wang, and al., " Evaluating the permutation and diffusion operations used in image encryption based on chaotic maps," Optik, vol. 127(2016), pp. 3541–3545.

[9] B. Norouzi and S. Murzakuchaki, "Breaking an image encryption algorithm based on the new substitution stage with chaotic functions," Optik, vol. 127 (2016), pp. 5695–5701.

[10] B. Wang, X. Wei, and Q.Zhang "Cryptanalysis of an image cryptosystem based on logistic map," Optik, vol. 124 (2013), pp. 1773–1776.

[11] F. Ozkaynak, A.B. Ozer "Cryptanalysis of a new image encryption algorithm based on chaos," Optik, vol. 127 (2016), pp. 5190–5192.

[12] Y. Dou, X.Liu and M.Li, "Cryptanalysis of a DNA and chaos based image encryption algorithm," Optik, vol. 145 (2017), pp. 456–464.

[13] H. Wang, D.Xiao, X.Chen and H.Hiang, "Cryptanalysis and enhancements of image encryption using combination of the 1D chaotic map," Signal processing, vol. 144 (2018), pp. 444–452.

[14] S.Dhall, S.K..Pal, and K.Sharma, "Cryptanalysis of image encryption scheme based on a new 1D chaotic system," Signal processing, vol. 146 (2018), pp. 22–32.

[15] F-G.Jeng, W-L.Huang, and T-H.Chen, "Cryptanalysis and improvement of two hyper-chaos-based image encryption schemes," Signal processing: Image communication, vol. 34 (2015), pp. 45–51.

[16] M.Li, Y.Guo, J.Huang and Y-Li, "Cryptanalysis of a chaotic image encryption scheme based on permutation-diffusion structure," Signal processing: Image communication, vol. 62 (2018), pp. 164–172.

[17] H. Zhou, X. Ling, " Problems with the chaotic inverse system encryption approach, " IEEE Trans. Circ. Syst. Vol. 44 (1997), pp. 268–271.

[18] G. Alvarez, S.Li, "Some basic cryptographic requirements for chaos based cryptosystems," Int. J. Bifurcation Chaos vol. 16 (2006), pp. 2129–2151.

An efficient model for management of road traffic in El-Oued city

*

Brahim Lejdel

University of El-Oued
El-Oued, Algeria
lejdel.brahim@gmail.com

***Abstract*—In the city, the road traffic is very studied in the last decade. It is also the most challenging topic. In this paper, we will propose an efficient model of the road network which is based on the Multi-Agent System and the Genetic Algorithms (MAS-GA). This model allows managing the road network in real time and permitting to regulate the traffic of vehicles in the intersection. Thus, we propose an optimal model that permits to synchronize the different components of road traffic as the vehicles, roads and traffic light to avoid the congestions that can occur principally in the crossroads.**

***Index Terms*—Multi-agent systems; Genetic Algorithms; Optimization;traffic control.**

I. Introduction

In the last decade, the use of mobility in the city is increased as vehicles, cars, and trains. However, this increasing caused many problems like congestion in road traffic. Thus, the role of this type of transport in urban areas has become an ever more important part of city life [1]. Economic growth and a modern lifestyle make inhabitants travel more frequently and for longer distances. Accordingly, the pressure for efficient and sustainable transport leads cities to invest in new transport technology and management of urban traffic [2]. Many city governments now use real-time analytics to manage aspects of how a city functions and is regulated [3]. The most common example relates to the movement of vehicles around a transportation network, where data from a network of cameras and sensors which can be set in the border of the road or on the buildings. These Data can be controlled by a system which permits to adjust traffic light sequences and speed limits and to automatically administer penalties for traffic violations [4]. We assume that the traffic congestion is a major factor, which is affecting any dense city. According to the direction of Transports in Algeria, a traffic flow increase in Algeria to 92% between 1988 and 2009. In addition, this trend is expected to continue for the near future, and the number of vehicles exceeds 8 million at the end of 2014. In addition, the accident rate in the country increased to more than 75%, in the same period. After the deep study of the subject, we find that we have three factors that can affect the congestion, as the traffic light phasings, the duration of each traffic lights and the interaction between the different actors of traffic as road, vehiclesetc. Traffic light phasings can affect traffic congestion in two different ways:

1) Traffic light phases at specific intersections are not optimized for the current traffic flow through that intersection.
2) Traffic light sequences at neighboring intersections are not synchronized correctly, if at all.

This paper is organized as the following. Firstly, we will present the related work in section 2. Then, section 3 describes our proposed approach which is based on two approaches, the Multi-Agent System and Genetic Algorithms (MAS-GA). In section 4, we describe the architecture of the system. Our experimental setup is presented in Section 5, where we describe the real-world case study that motivated this research. Finally, we add a conclusion and some future works.

II. related works

Traffic congestion is a term used when many vehicles are clogged in one place and there is very slow or no movement. Thus, we can define the traffic congestion as the way in which vehicles interact to impede each other progress. These interactions and their influence on journeys usually increase as demand for the available road space approaches capacity or when capacity itself is reduced through road works or closures for example. In addition, other events such as bad weather or road traffic accidents can also have a significant influence on congestion [5]. The real-time control of traffic-lights [6] is not feasible because of various reasons (legal, technical, etc.), and we must instead find a highly-reliable global schedule of traffic-lights that works well in the dynamic and uncertain traffic system [7]. For optimizing the light cycle programs of traffic signals within a city in order to improve traffic flow and reduce congestion, there are many works that use the multi-agent system to model an intelligent transportation system [8], [9]. In these works, each vehicle has a cooperative behavior with the intersection and communicates its trajectory

when it enters the approach of the intersection. With its computation abilities and the available information, a vehicle runs a solver to produce the optimal conguration according to its currents situation. Several approaches have focused on the traffic lights in order to minimize the delays and queue lengths [10]. There are many works which use the fitness for calculating the reliability of a candidate traffic-light program is evaluated by simulating vehicle routes and velocities over a given traffic network [10], [11]. The simulation of traffic flows on a specific city requires collecting network data (topology of the area and information about traffic lights), which is usually precise and static, and traffic data (number of vehicles, their journeys, and velocities), which is estimated from highly dynamic real-world data. Appropriate sensor placement is a fundamental requirement in the control of any intersection [12], especially when traffic light recognition is needed for autonomous vehicles [19]. Despite the inherent uncertainty of simulated traffic scenarios, the literature on traffic-light optimization often relies on deterministic simulation on a single traffic scenario [10], [11], [13]. When multiple scenarios are considered, they are actually used for evaluating the flexibility of the optimization algorithm by optimizing each scenario separately [14]. More recently, Ferrer et al. [15] validated the reliability of candidate solutions on multiple scenarios after the optimization phase, however, each solution is still optimized with respect to a single scenario. In this paper, we will propose a hybrid approach by combining multi-agent systems with Genetic Algorithms.

III. Proposed Approach

For proposing a complete model that can resolve the traffic problem, we have to define certain concepts.

1) The Multi-agent system (MAS): An agent is a software system that is situated in some environment, and that is capable of autonomous action in order to meet its design objectives [16].
2) The Genetic Algorithm: Genetic algorithms are developed in [17] to imitate the phenomena adaptation of living beings. They are optimization techniques based on the concepts of natural selection and genetics. It searches an optimal solution among a large number of candidate solutions within a reasonable time (the process of evolution takes place in parallel) [18].

A. Intersection Agents

The Intersection Agent has been designed with three components. The components are:

1) An optimizer in this case a genetic algorithm.
2) A traffic flow simulator.

B. Negotiation between traffic factors

Negotiation techniques have been used in multi-agent systems for several purposes, such as military negotiation, auction, resource allocation, task allocation, transportation, and conflict resolution [19], [20]. For example, in order to solve congestion during the traffic, the vehicle agents negotiate, each one trying to obtain enough space wherein to find the best place for the vehicle in the road according to their directions and their current positions. Therefore, when conflicts occur between vehicles, it is important to limit their effects. In Figure 1, we present an example of negotiation between two vehicle agents; vehicle Agent 1 and vehicle Agent 2. Thus, these two agents negotiate by proposing a plan of actions according to the direction of the vehicle and its current position.

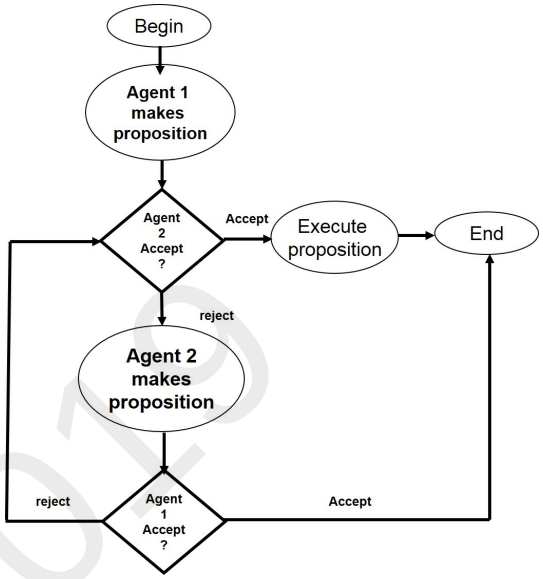
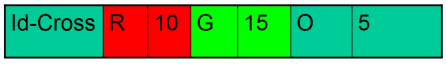


Fig. 1. Negotiation between two agents.

C. Representation of Gene

Figure 2 presents the structure of the gene.



Id-Cross: intersection identify
R : red, G: green, O: Orange

Fig. 2. Gene structure of intersection agent.

IV. System Architecture

Figure 3 shows the system architecture of our system.

The detailed architecture of the Intersection agent leads to a simple but effective architecture, such as shown in Figure 4. Intersection Agents manage traffic flow through their intersections and pass two types of traffic flow data to each other:

1) Input data: information on how many vehicles can enter from the agents neighbor before its entry roads become congested.
2) Output data: information on how many vehicles are expected to exit one agent and enter its neighbor.

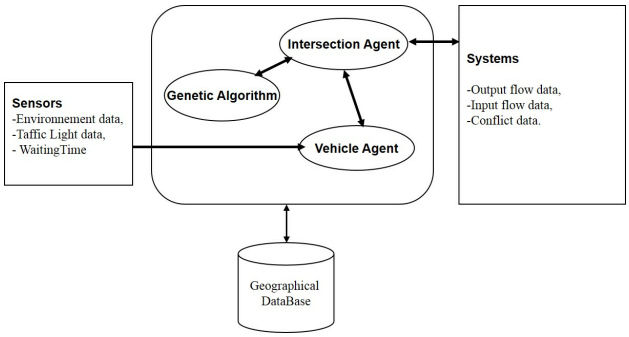


Fig. 3. The architecture of our system.

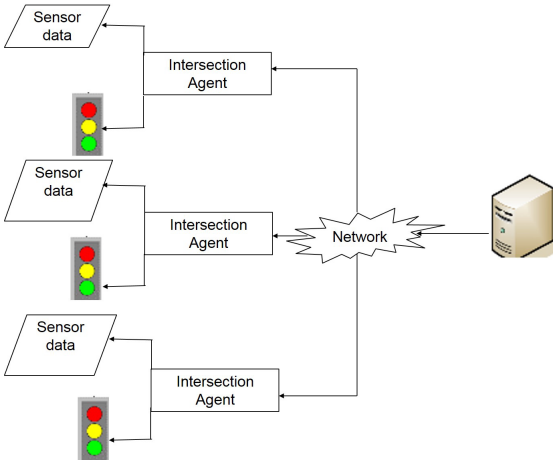


Fig. 4. The detailed architecture of of Intersection agent.

V. Experimentation and Results

The use of a Genetic Algorithm for the optimizer has a major advantage over systems that rely on predefined sequences, as the Genetic Algorithm enables each Intersection Agent to discover sequences that may not resemble any predefined sequence, but may be optimal for the current traffic conditions. Figure 5 shows an example event sequence. Here, the intersection is represented by a letter and the time by a number.

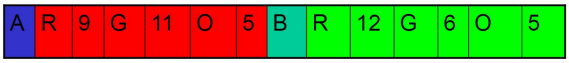


Fig. 5. An example an optimal solution.

A. Crossover and mutation

In this section, we define the crossover operator. It defines the procedure for generating a child from two parent chromosomes. The crossover operator produces new individuals as offspring, which share some features taken from each parent. Figure 6 shows the single point crossover. Next, the mutation operator is important. In this work, the results presented here were generated using a 1% mutation probability, which was determined experimentally, utilizing a single case of the vector of the traffic light. We present the random mutation in Figure 7.

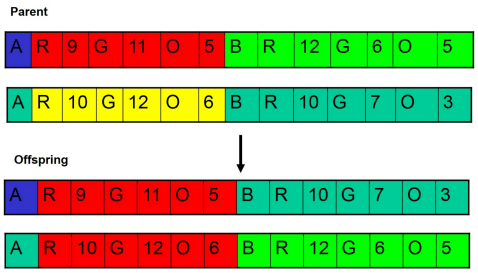


Fig. 6. Single point crossover

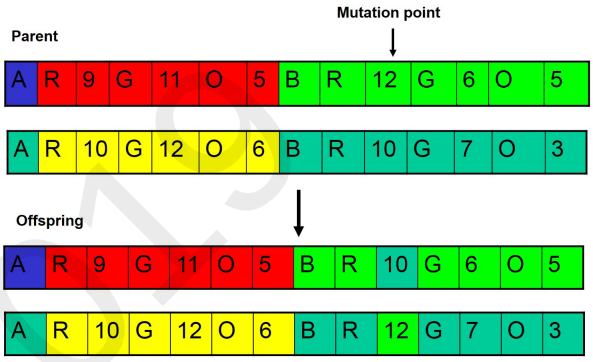


Fig. 7. Mutation operator.

B. Evaluation of solutions

The objective function used to evaluate solutions to the traffic flow requires a number of definitions that model the problems underlying structure, specifically:

- InterSec = A,B,C,.....is the set of intersection agents.
- Vhcle = V1,V2,V3,..... is the set of all vehicles in the intersection.
- N1 is the number of vehicles which enter the intersection.
- N2 is the number of vehicles which do not exit the intersection.
- T1 is the time when the vehicle entered the intersection.
- T2 is the time when the vehicle exited the intersection.

Thus, we are three different factors which used to evaluate the solutions. Input flow data, output flow data and waiting time of vehicles in the intersection. In this context, we have mainly two important functions $f(N)$ and $f(T)$ which permits evaluating the performance and efficiency of the proposed approach. These two functions are calculated by the intersection agent. Firstly, The objective function $f(N)$ control the total number of the vehicle in the intersection. Thus, we have in Eq.1.

$$f(N) = \sum_{k=1}^{n} (N1 - N2) \quad (1)$$

The second objective function $f(T)$ calculate the total waiting time in the intersection. Thus, we can define this objective function in the Eq.2.

$$f(T) = \sum_{k=1}^{n} (T2 - T1) \quad (2)$$

The objectives of this optimization mechanism are to maximize the number of vehicles $f(N)$ passed an intersection and to minimize the total waiting time $f(T)$ for evaluating the performance and the efficiency of our system. The optimizer can stop the execution of the genetic algorithm for three reasons, when it finds certain fitness, when it achieves a fixed number of iterations or when a certain time of execution is exceeded. In the latter two cases, we choose the solution that has the best fitness.

C. Simulation & Results

In this section, we present our implementation of Multi-agent system and the genetic algorithm for solving road traffic problem as congestion. We know that the maximum waiting time of a vehicle in the intersection is an important factor to measure the efficiency of such a system for traffic flow [21]. Thus, our system should try to minimize the waiting time for the vehicle over a number of generations. Firstly, we introduce the SUMO (Simulator of Urban Mobility) and the SUMO Cycle Program Generator (SCPG) algorithm provided with the SUMO package [22], [23]. To test the advantage of our approach, we use real data. We use Data from El-Oued Town because it knows great congestion. El-Oued Town computes three principal intersections which are known as the black point which are Choot intersection (Intersection 1), Station Intersection (Intersection 2) and Willaya Intersection (Intersection 3). The data used in this experimentation are based on the observation of traffic flow in three intersections in El-Oued Town. This observation is obtained in the peak-hour period as 8h00-10h00 and 17h00-18h00. Figure 8 present the road network used in this study. The three Intersection Agents communicate and negotiate with each other to ensure that traffic flow is optimized across the entire area and that an over-optimization of traffic flow through an intersection by one agent does not cause subsequent problems for its neighbors.

Tables I shows the number of vehicles per minute that traverse the three intersections mentioned above. Thus, we present the traffic flow data in such an intersection.

Intersection 1	**Vehicles per minute**
Tiksebt to 08 March	200
Hassani Abd Kerim to Mellah	150
Intersection 2	**Vehicles per minute**
Souk Libya to Mih Ounsa	120
Tiksebt to Place 35	220
Intersection 3	**Vehicles per minute**
Kouinine to Willaya Intersection	250
8 march to Mih Ounsa	160

TABLE I
VEHICLE RATES FOR INTERSECTION AGENTS.

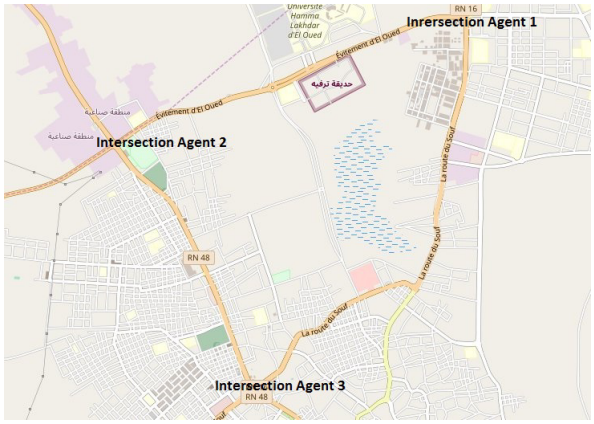


Fig. 8. Intersection Agents for El-Oued Town.

1) Study Case 1: The following experiments were set up:

- Experiment 1. We will use one Agent which control multiple intersections. In this experiment, one Intersection Agent was used to optimize the traffic light phase sequences for both intersections.
- Experiment 2. We will use multi-agents which control intersections. In this experiment, three Intersection Agents (one manage the intersection 1, the second manage intersection 2 and other manage intersection 3) were used to evaluate the performance of collaborative Intersection Agents. For this experiment, each Intersection Agent was run on a separate PC.

The experiments were run on a cluster of 4 machines with Intel Core2 Quad processors Q9400 (4 cores per processor) at 2.66GHz and 4GB memory. The cluster was managed by HTCondor 8.2.7, which allowed us to perform parallel independent executions to reduce the overall experimentation time. We can see clearly that by using multiple Intersection Agents, traffic flow through the road network is optimized in fewer generations, and in significantly less time, than when using one agent for multiple intersections. After 300 generations, the waiting time achieved by Experiment is 120 seconds, whereas the waiting time achieved by Experiment 2 was 17 seconds. Thus, Intersection Agents in Experiment 2 converge to their optimal solution at a significantly faster rate than the all-in-one solution from Experiment 1, as shown in Figure 9.

2) Study Case 2: In this study case, we compare the result obtained by Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The results obtained by PSO remain significantly worse than the ones obtained by GA. This difference in solution fitness between PSO and the GA is due to the slow convergence of PSO, as shown by the plot of fitness over a number of generations in Figure 10. Each line in the plot is the fitness value of the best solution, as estimated by each algorithm at each moment of its execution, averaged over all independent repetitions for each algorithm. PSO is a population-based meta-heuristic with many different variants.

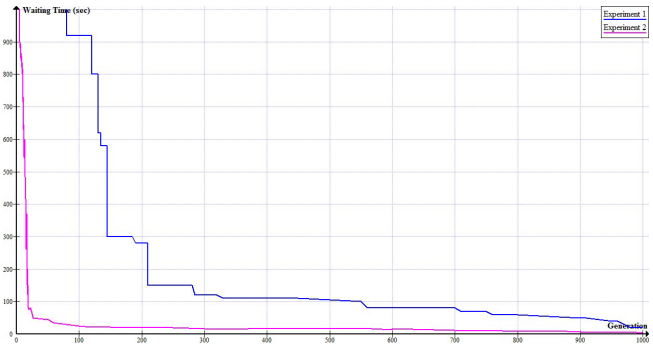


Fig. 9. Waiting time over generations.

We use here the Standard PSO introduced in [24],

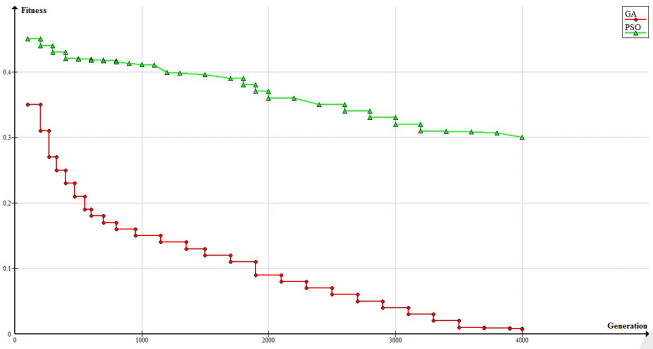


Fig. 10. Convergence of objective function for three Agent intersections.

VI. Conclusion and Future Works

The objective of traffic flow optimization consists to propose an efficient and optimal model to avoid congestion in the transportation system over the city roads. In this paper, we defined and presented a hybrid approach which permits to optimize the traffic flow. This approach is based; on one hand on the Multi-agent system, and on the other hand on the genetic algorithm. This approach is used to increase the efficiently of the traffic flow and avoid conflicts which increase the waiting time of vehicles in intersections. In this work, we use the simulation to test our approach in three intersections. This simulation presents an interconnection of the different traffic actors such as roads, vehicles, and sensors that provide data on the traffic flow in real-time. To perform this simulation we create a geographical database of the road network. The initial results of the experimentation are very good which permit to encourage us to continue in this area of research. In future work, we want to compare our obtained solutions with the one currently used on the real system. Also, we want to study the results of simulation optimization algorithms for the traffic light under unplanned and unexpected changes to traffic flow, such as a road closure due to an accident.

References

[1] Martin, D. Rob, K, The automatic management of drivers and driving spaces,, Geoforum, 38, 2007, 264275.

[2] Marie T., Dinesh M., Geetam T., Sustainable transport and the modernization of urban transport in Delhi and Stockholm,Cities, 27, 2010, 421429.

[3] Rob K., The real-time city? Big data and smart urbanism, GeoJournal, 79, 2014, 114.

[4] Dodge, M. and Kitchin, R., The automatic management of drivers and driving spaces, Geoforum 38(2), 2007, 264275.

[5] John, S., Peter, N., Maarten V., Urban traffic incident management in a digital society: An actor-network approach in information technology use in urban Europe, Technological Forecasting and Social Change, 89, 2014, 245-261.

[6] Z. Cao, S. Jiang, J. Zhang, H. Guo, A unified framework for vehicle rerouting and traffic light control to reduce traffic congestion, IEEE Trans. Intell.Transp. Syst., 18 (7), 2017, 19581973.

[7] Y. Bravo, J. Ferrer, G.J. Luque, E. Alba, Smart mobility by optimizing the traffic lights: A new tool for traffic control centers, in: E. Alba, F.Chicano, G. Luque (Eds.), Smart Cities, Lecture Notes in Computer Science, Springer, 2016, 147156.

[8] S. Wang, N.U. Ahmed, T.H. Yeap, Yeap, Optimum management of urban traffic flow based on a stochastic dynamic model, IEEE Trans. Intell. Transp. Syst, 2018, 113.

[9] A.A. Juan, J. Faulin, S.E. Grasman, M. Rabe, G. Figueira, A review of simheuristics: Extending metaheuristics to deal with stochastic combinatorial optimization problems, Oper. Res. Perspect., 2, 2015, 6272.

[10] J. Garca-Nieto, E. Alba, A.C. Olivera, Swarm intelligence for traffic light scheduling: Application to real urban areas, Eng. Appl. Artif. Intell, 25 (2), 2012, 274283.

[11] Z. Li, M. Shahidehpour, S. Bahramirad, A. Khodaei, Optimizing traffic signal settings in smart cities, IEEE Trans. Smart Grid, 8(5), 2016, 2382-2393.

[12] L.C. Zammit, S.G. Fabri, K. Scerri, Scerri, Real-time parametric modeling and estimation of urban traffic junctions, IEEE Trans. Intell. Transp. Syst., 2019, 111.

[13] M. Pres, G. Ruiz, S. Nesmachnow, A.C. Olivera, Multiobjective evolutionary optimization of traffic flow and pollution in Montevideo, Uruguay, Appl.Soft Comput. J., 2018.

[14] J. Snchez, M. Galn, E. Rubio, Applying a traffic lights evolutionary optimization technique to a real case: Las Ramblas area in Santa Cruz de Tenerife, IEEE Trans. Evol. Comput., 12 (1), 2008, 2540.

[15] J. Ferrer, J. Garca-Nieto, E. Alba, F. Chicano, Intelligent testing of traffic light programs: Validation in smart mobility scenarios, Math. Probl. Eng., 2016, 119.

[16] Vitor L., Isabel A., Erik D., Jean-Michel G., Smart Cities initiative: How to foster a quick transition towards local sustainable energy systems, Final Report, THINK, 2011.

[17] DeJong K. and Sarma J., Generation Gaps Revisited, Foundations of Genetic Algorithms, D. Whitley, Morgan-Kaufmann Publishers, San Mateo, 1993.

[18] Holland J., Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Harbor, 1975.

[19] Gradinescu, V. Gorgorin, C., Diaconescu, R., Cristea, V.,Iftode, L., Adaptive traffic lights using car-to-car communication, Vehicular Technology Conf, IEEE 65th, 66, 1993, 2125.

[20] Zhang, X., Lesser, V., Multi-Linked Negotiation in Multi-Agent System, Proceedings of the First International Joint Conference on Autonomous Agents And Multi-agent Systems, AAMAS 2002, 2002, 1207-1214.

[21] Victor G., Cristian G., Liviu I., Adaptive Traffic Lights Using Car-to-Car Communication, IEEE, 66, 2007, 21-25.

[22] M. Behrisch, L. Bieker, J. Erdmann, D. Krajzewicz, SUMO - simulation of urban mobility: An overview, in: SIMUL 2011,The Third International Conference on Advances in System Simulation, Barcelona, Spain, 66, 2011, 6368.

[23] D. Krajzewicz, J. Erdmann, M. Behrisch, L. Bieker, Recent development and applications of SUMO - simulation of urban mobility, Int. J. Adv. Syst. Meas, 5(34), 2012, 128138.

[24] M. Clerc, J. Kennedy, Standard PSO, Particle Swarm Central, http://www.particleswarm.info/, 2011.

MÉTHODE D'OPTIMISATION DU SOUS-GRADIENT ET APPLICATIONS

RACHID BELGACEM[1] AND AHMED BOKHARI[2]

Résumé. Ce travail est consacré aux méthodes d'optimisation de sous-gradient qui sont utilisées pour résoudre des problèmes de programmation linéaire en nombres entiers, en utilisant la méthodologie de la relaxation Lagrangienne. En pratique, les problèmes de programmation linéaire en nombres entiers pertinents sont généralement composés de quelques structures mathématiques agréables jointes à certaines conditions compliquées, ce qui rend toute procédure de résolution difficile. Différentes variantes de méthodes de sous-gradient et fréquemment utilisé pour résoudre un tel problèmes. Par la suite, des applications numériques de ces méthodes au problème TSP sont tenues pour faire une comparaison en terme de la borne inférieure du tour optimal.

1 Introduction

Ce travail est consacré aux méthodes d'optimisation de sous-gradient qui sont utilisées pour résoudre des de programmation linéaire en nombres entiers, en utilisant la relaxation Lagrangienne. Ceci implique le problème de la maximisation d'une fonction duale Lagrangienne concave mais non différentiable. De nombreux problèmes de programmation linéaire en nombres entiers d'intérêt pratique sont de grande taille comme le problème du voyageur de commerce, appelé aussi (en anglais "Travelling Salesman Problem") noté TSP. Dans ce problème : Soit un représentant de commerce qui doit visiter un certain nombre de clients situés dans n villes. Le coût de déplacement d'une ville i à une ville j est donné. L'objectif du représentant de commerce est de parcourir toutes les villes avec un coût total qui soit minimum, mathématiquement, cela consiste à déterminer, dans un graphe, un cycle Hamiltonien qui soit de longueur minimum.

En pratique, les problèmes de programmation linéaire en nombres entiers pertinents sont généralement composés de quelques structures mathématiques agréables jointes à certaines conditions compliquées (ou contraintes), ce qui rend toute procédure de résolution difficile. La méthode de relaxation Lagrangienne soulève ces contraintes compliquées et utilise la structure spéciale pour résoudre le problème relaxé. Cela produit une borne inférieure pour un problème de minimisation.

Un défi majeur dans la méthode de relaxation lagrangienne d'une minimisation d'un problème de programmation linéaire en nombres entiers est de maximiser efficacement la fonction duale lagrangienne qui n'est définie que de façon implicite, n'est pas différentiable, concave et affine par morceaux.

La méthode de sous-gradient est fréquemment utilisée pour résoudre de tels problèmes puisqu'elle n'exige pas de différenciation. En fait, c'est une procédure itérative qui recherche une solution optimale en utilisant la direction du vecteur gradient à chaque point où le gradient de la fonction existe. Mais, à un point où le gradient n'existe pas, on remplace le vecteur gradient par un vecteur sous-gradient. Le sous-gradient d'une fonction à un point n'est généralement pas unique.

Différentes variantes de méthodes de sous-gradient seront étudiées et une présentation unifiée des méthodes sera donnée. L'inconvénient central de ces méthodes est que leur convergence est généralement lente qui est principalement causée par ce qu'on appelle les phénomènes de zigzag. La lenteur de la convergence est un problème crucial, en particulier, compte tenu de la nécessité de résoudre des problèmes à grande échelle. En effet, différentes versions de méthodes de sous-gradient telles que la stratégie de direction moyenne, la technique du gradient modifié et une nouvelle méthode de sous-gradient dévié modifié serons développées dans le but d'améliorer la

vitesse de convergence de la méthode sous-gradient pur en contrôlant ses phénomènes de zigzag. Par la suite, des applications numériques de ces méthodes au problème TSP sont tenues pour faire une comparaison en terme de la borne inférieure du tour optimal.

A la fin, nous terminerons notre travail par une conclusion.

2 Relaxation et dualité Lagrangienne

Considérons le programme linéaire en nombre entier suivant :

$$(ILP)\begin{cases} \min\ z = c^T x \\ Ax \geq b \\ x \in \mathbb{X} = \{x \in \mathbb{Z}_+^n : Dx \geq d\} \end{cases} \tag{2.1}$$

Avec (A, b) et (D, d) sont des matrices $m \times (n+1)$ et $r \times (n+1)$ respectivement, $c \in \mathbb{R}^n$ et $x \in \mathbb{Z}_+^n$. Nous appelons le problème (ILP) le problème primal et sa solution une solution primale.

La relaxation Lagrangienne est appliquée en générale lorsqu'on reconnaît des contraintes difficiles dont la relaxation engendrera un problème plus facile à résoudre. La relaxation Lagrangienne s'articule au tour de l'idée qui consiste à relâcher les contraintes difficiles ,non pas en les supprimant, mais en les prenant en compte dans la fonction objectif de sorte qu'elle pénalisent la valeur des solutions qui valent ces dernière. Supposons que les contraintes $Ax \geq B$ sont "difficiles" dans le sens où on suppose que l'on dispose d'un algorithme efficace pour minimiser la fonction z sur l'ensemble X. les contraintes relachées sont réinjectées dans la fonction objectif, pondérées par les coefficients $u \in \mathbb{R}_+^m$ appellés multiplicateurs de Lagrange. La fonction Lagrangienne est donnée par :

$$L(x, u) = c^T x + u^T(b - Ax)$$

Soit $u \in \mathbb{R}_+^m$, le problème de la relaxation Lagrangienne est défini comme suit :

$$(L_u)\begin{cases} \Phi(u) = \min\ L(x, u) \\ x \in \mathbb{X} \end{cases} \tag{2.2}$$

La fonction Φ est appelée "fonction duale" et les coefficients $u \in \mathbb{R}_+^m$ sont appelés variables duales.

On définit le problème dual Lagrangien du problème (ILP) comme suit :

$$(LD)\begin{cases} \Phi^* = \max\ \Phi(u) \\ \quad u \in \mathbb{R}_+^m. \end{cases} \tag{2.3}$$

Lorsque les m contraintes qui sont dualisée sont des contraintes d'égalité de la forme "$Ax = b$", les multiplicateurs de Lagrange correspondant sont de signe quelconque $u \in \mathbb{R}^m$.

Theorem 2.1. ***(Théorème de la dualité faible) :*** *Soient les problèmes (ILP), (L_u) et (LD) définis par (2.1),(2.2) et (2.3) respectivement et x une solution réalisable de (ILP). Alors pour tout $u \geq 0$,*

$$\Phi(u) \leq \Phi^* \leq z^*$$

Theorem 2.2. ***(Concavité de la fonction duale) :***
La fonction duale $u \to \Phi(u)$ est une fonction concave linéaire par morceaux.

3 Méthodes sous-gradient en optimisation

La valeur maximale d'une fonction différentiables concave peut être généralement déterminée par les méthodes du gradient. Une méthode du gradient, par exemple la méthode de la plus grande pente, trouve une solution optimale du problème $\max_x f(x)$ par méthode itérative dans laquelle, à partir de x^0, une suite de x^n convergeant finalement vers une solution optimale est construite selon la relation :

$$x^{n+1} = x^k + \lambda_n \nabla f(x^n)$$

où $\lambda_n \geq 0$ est un pas approprié et $\nabla f(x^n)$ est le vecteur gradient de f en x^n. Dans le cas de notre problème, la fonction duale n'est pas différentiable. Donc, on ne peut pas utiliser la

méthode du gradient car il y a des points où $\nabla\Phi$ n'existe pas. Dans ce cas, les gradients sont remplacés par des sous-gradients afin d'utiliser la structure de la concavité de la fonction duale.

Definition 3.1. (sous-gradient) : Soit $f : \mathbb{R}^m \to \mathbb{R}$ une fonction concave. Un vecteur $s \in \mathbb{R}^m$ est appelé sous-gradient de f au point $\bar{x} \in \mathbb{R}^m$ si

$$f(x) \leq f(\bar{x}) + s^T (x - \bar{x}) \qquad \forall x \in \mathbb{R}^m \tag{3.1}$$

Definition 3.2. (sous-différentiel) : Le sous-différentiel de f au point $\bar{x}$ est l'ensemble de touts les sous-gradients de f au point $\bar{x}$ est donné par :

$$\partial f(\bar{x}) = \left\{ s : f(x) \leq f(\bar{x}) + s^T (x - \bar{x}) \quad \forall x \in \mathbb{R}^m \right\}.$$

Theorem 3.3. *Soit $f : \mathbb{R}^m \to \mathbb{R}$ concave et différentiable. Alors, $\nabla f(\bar{x}) \in \partial f(\bar{x}) \quad \forall x \in \mathbb{R}^m$.*

Notons que, la définition de la concavité signifie qu'un sous gradient est un vecteur de l'hyperplan supportant l'épigraphe de f en $(\bar{x}, f(\bar{x}))$, qui est un ensemble convexe fermé.

Theorem 3.4. *Si $\nabla f(x)$ existe, alors $\partial f(x)$ est un singleton et $\partial f(x) = \{\nabla f(x)\}$.*

Theorem 3.5. *Le sous-différentiel $\partial f(\bar{x})$ de f en $\bar{x} \in \mathbb{R}^m$ est un ensemble convexe.*

Theorem 3.6. *Une condition nécessaire et suffisante pour que $x^* \in \mathbb{R}^m$ soit un maximum d'une fonction concave f sur $\mathbb{R}^m$ est $0 \in \partial f(x^*)$.*

Notons que la condition "$0 \in \partial f(x^*)$" est une généralisation de la condition stationnaire usuelle "$\nabla f(x^*) = 0$" dans le cas différentiable.

Theorem 3.7. *Soit x_u une solution optimale du problème (L_u) . Alors : $s = b - Ax$ est un sous-gradient de Φ au point u.*

Le schéma standard de la méthode sous-gradient de base pour résoudre (LD) est le suivant :
On commence par un certain point u^0 et on construit une séquence de points $\{u^k\}$ selon la règle

$$u^{k+1} = P^+ \left(u^k + \lambda_k s^k \right), \quad k = 0, 1, 2, ... \tag{3.2}$$

où s^k est le sous gradient de Φ au point u^k, λ_k est le pas, et P^+ est l'opérateur projection, de $\mathbb{R}^m$ sur $\mathbb{R}^m_+$ c-à-d :

$$P^+(u) = \max(0, u) = (\max(0, u_1), ..., \max(0, u_m))^\top$$

À un point donné, nous n'avons aucun sous-gradient unique de la fonction. Cela pose certaines difficultés en ce qui concerne la construction d'une bonne procédure itérative qui utilise un vecteur de sous gradient comme direction de progression.

En général dans le cas différentiable, il est bien connu que le vecteur gradient est un vecteur de descente, ce n'est pas le cas pour un vecteur de sous-gradient et donc la procédure d'itération de la méthode de sous-gradient n'améliore pas nécessairement la valeur de la fonction objectif à certaines étapes.

4 Algorithme sous-gradient dévié

Un autre défi dans l'optimisation de sous-gradient est le choix de la direction de recherche qui affecte la performance de calcul de l'algorithme. On sait que le choix de la direction du sous-gradient s^k conduit au phénomène de zig-zag qui pourrait ralentir la procédure d'exploration vers l'optimalité. Pour surmonter cette situation, dans l'esprit de la méthode du gradient conjugué [J. Nocedal (1999), R. Fletcher (1964)], nous pouvons adopter une direction de recherche qui dévie la direction du sous-gradient pure. Par conséquent, la direction de recherche d^k à u^k est calculée comme suit :

$$d^k = s^k + \delta_k d^{k-1}, \tag{4.1}$$

où $\delta_k \geq 0$ est un paramètre de déviation, s^k est un sous-gradient de la fonction Φ à u^k et d^{k-1} est la direction précédente ($d^0 = 0$). Ensuite, la nouvelle itération est calculée comme suit :

$$u^{k+1} = P_\Omega \left(u^k + \lambda_k d^k \right). \tag{4.2}$$

Definition 4.1. Soit u_k un scalaire positif et $d^k \in \mathbb{R}^n$. On dit que la procédure

$$u^{k+1} = P_X(u^k + \lambda_k d^k), k = 0, 1, 2, ...$$

forme un **zig-zag**, si l'angle entre les directions d^{k+1} et d^k est obtus, i.e., $d^{k+1}d^k < 0$.

L'algorithme suivant donne une description détaillée de la méthode du sous gradient dévié :

(1) Choisir un vecteur initiale u^1, et soit $k = 1$, $d^k = 0$.

(2) Déterminer un sous-gradient $s^k \in \partial\Phi(u^k)$

$$d^k = s^k + \delta_k d^{k-1} \tag{4.3}$$

$$u^{k+1} = P^+(u^k + \lambda_k d^k) \tag{4.4}$$

Les règles pour déterminer δ_k et t_k seront données plus tard.

Répéter la procédure à partir de l'étape 2 tant que le test d'arrêt n'est pas vérifie.

(3) Si une condition d'arrêt n'est pas encore prise,on revient à l'étape 2.$\hookrightarrow Test$

Un comportement important de la procédure du sous-gradient est que, à chaque itération, la direction du sous gradient s^k forme un angle aigu avec $(u^* - u^k)$.

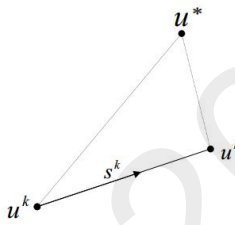


FIGURE 1. s^k forme un angle aigu avec $(u^* - u^k)$.

Toutefois, selon plusieurs travaux (voir par exemple, [P.M. Camerini(1975)], [H.D. Sherali(1989)] et [R. Belgacem (2017)], l'angle entre la direction du sous-gradient s^k peut former un angle obtus avec la direction précédente s^{k-1}. Ce qui peut forcer le prochain itéré de devenir proche du précédent. Ce phénomène peut évidemment ralentir la convergence de la procédure.

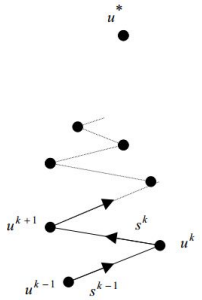


FIGURE 2. Phénomène de zig-zag

De tel phénomène de **zig-zag** pourrait se manifester à n'importe quelle étape de la procédure du sous-gradient et ralentir le processus de recherche.

Afin d'éviter un tel comportement, P.M. Camerini et al [P.M. Camerini(1975)] ont proposé une modification de la méthode de sous-gradient pur dans laquelle la direction du sous-gradient s^k à l'itération k est remplacé par une direction sous-gradient dévié d^k_{MGT}, donnée par :

$$d^k_{MGT} = s^k + \delta_k^{MGT} d^{k-1}. \tag{4.5}$$

Avec le paramètre de déviation δ_k^{MGT} est donné par :

$$\delta_k^{MGT} = \begin{cases} -\eta_k \frac{s^k d^{k-1}}{\|d^{k-1}\|^2} & \text{if } s^k d^{k-1} < 0, \\ 0 & \text{sinon,} \end{cases} \tag{4.6}$$

où $0 < \eta_k \leq 2$ et $d_{MGT}^{k-1} = 0$ pour $k = 0$.

Cette stratégie, est appelée la technique de gradient modifié (Modified Gradient Technique "MGT").

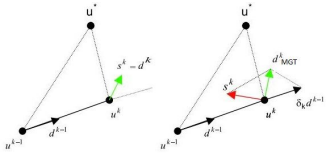


FIGURE 3. Illustration dans un cas bidimensionnel : cas de Camerini.

Il existe diverses formes du choix du paramètre de déviation différentes de celle proposée par Camerini et al, H.D. Sherali et al [H.D. Sherali(1989)] ont proposé de choisir la direction de déviation comme la bissectrice de l'angle formé par le sous-gradient actuel s^k et la direction précédente. Pour obtenir cette direction, le paramètre de déviation est calculé selon

$$\delta_k^{ADS} = \frac{\|s^k\|}{\|d^{k-1}\|}. \tag{4.7}$$

Avec ce choix du paramètre de déviation δ_k^{ADS}, la direction devient :

$$d_{ADS}^k = s^k + \delta_k^{ADS}\, d^{k-1}. \tag{4.8}$$

On appelle cette stratégie direction moyenne(Average Direction Strategy "ADS").

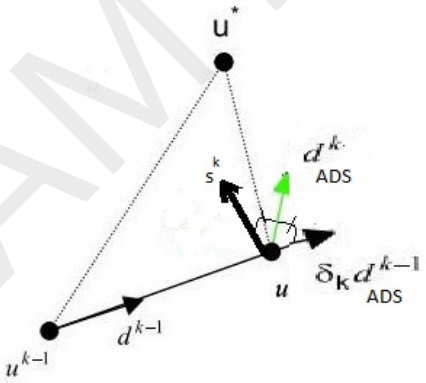


FIGURE 4. Illustration dans un cas bidimensionnel : cas de Sherali.

Autrement, une nouvelle méthode de sous-gradient dévié modifié (a New Modified Deflected Subgradient method "NMDS") [?] qui détermine la direction de recherche déviée comme une combinaison convexe de la direction d_{MGT}^k (4.5) et de la direction d_{ADS}^k (4.8). Cette nouvelle direction est définie comme suit :

$$d_{NMDS}^k = (1 - \alpha_k) d_{MGT}^k + \alpha_k d_{ADS}^k \ , \quad \alpha_k \in (0, 1)\,. \tag{4.9}$$

On obtient alors le paramètre de déviation suivant :

$$\delta_k = \begin{cases} \frac{-\eta_k (1-\alpha_k) s^k d^{k-1} + \alpha_k \|s^k\| \|d^{k-1}\|}{\|d^{k-1}\|^2} & \text{si } \ s^k d^{k-1} < 0, \\ 0 & \text{sinon,} \end{cases} \tag{4.10}$$

où $\alpha_k = -\cos(s^k, d^{k-1})$ si $s^k d^{k-1}$. Avec $0 < \eta_k \leq \frac{1}{2-\alpha_k}$.
Par conséquent $d^k_{NMDS} = s^k + \delta_k d^{k-1}$.

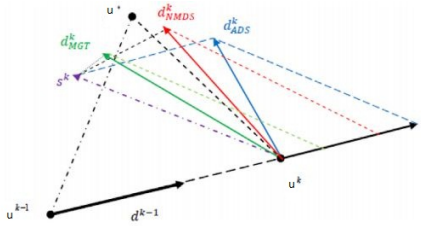


FIGURE 5. Cas où s^k est dévié car il a formé un angle obtus avec d^{k-1} et la direction d^k_{NMDS} est meilleure par rapport aux autres directions d^k_{ADS} et d^k_{MGT}.

5 Résultats numériques

Pour une comparaison numériques entre les quatre algorithmes : pur, MGT, ADS et NMDS, nous avons opté le problème de voyageur de commerce symétrique ($TSPs$) . Il peut être formulé comme suit :

$$\min \quad \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} c_{ij} x_{ij}, \tag{5.1}$$

sous contraintes

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1, ..., n, \tag{5.2}$$

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = 1, ..., n, \tag{5.3}$$

$$\sum_{i\in S}^{n} \sum_{j\in S}^{n} x_{ij} \leq |S| - 1, \quad \forall S : 2 \leq |S| \leq n-2, \tag{5.4}$$

$$x_{ij} = 0 \text{ or } 1, \quad i, j = 1, ..., n, \tag{5.5}$$

où les c_{ij} sont les coûts du lien (i, j) et $S \subset \{1, ..., m\}$. Si X est l'ensemble des 1-arbres [?], les contraintes de sous-tour (5.4) peuvent être éliminées en insistant sur le fait qu'un vecteur x satisfaisant les contraintes (5.2), (5.3) et (5.5) doit aussi appartenir à X.

En particulier pour le cas symétrique, les contraintes (5.2) et (5.3) peuvent être remplacées par (5.6) conduisant à la formulation équivalente suivante de TSPs :

$$\min \quad \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} c_{ij} x_{ij},$$

sous contraintes

$$\sum_{\substack{j=1 \\ j\neq i}}^{n} x_{ij} + \sum_{\substack{j=1 \\ j\neq i}}^{n} x_{ji} = 2 \quad \text{for } i = 1, ..., n, \tag{5.6}$$

$$x \in X.$$

De là, on obtient la fonction duale suivante, qui doit être maximisé :

$$\Phi(u) = \min \left\{ \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j\neq i}}^{n} (c_{ij} + u_i + u_j) x_{ij}, \ x \in X \right\} - 2 \sum_{i=1}^{n} u_i, \tag{5.7}$$

où $u \in \mathbb{R}^n$ est le vecteur des multiplicateurs de Lagrange.

Étant donné un vecteur $\overline{u}$, si $\overline{x}$ optimise $\Phi(\overline{u})$, alors un vecteur $\overline{s}$ dont la i^{me} composante.

$$\overline{s_i} = \left(\sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ij} + \sum_{\substack{j=1 \\ j \neq i}}^{n} x_{ji} - 2 \right) \tag{5.8}$$

est un sous-gradient de $\Phi(u)$ à $\overline{u}$ ([M.S. Bazaraa (2006)] and [M.H. Held(1974)]).

Les quatre algorithmes ont été implémentés par Matlab et exécutés sur un processeur Intel (R) Core (TM)2 Duo T7300 @ 2.00GHz 2.00GHz RAM 2.00GO.

Le tableau 6 montre les résultats expérimentaux obtenus par : Algorithme pur, stratégie MGT, stratégie ADS et algorithme NMDS avec 10 tests des instances symétriques pris à partir de TSPLIB. Cependant, toujours l'algorithme NMDS surperforme les autres en nombre d'itérations et en temps d'exécution. Ce tableau montre également que la stratégie NMDS donne des résultas quasi-optimaux pour plusieurs instances (voir figures)Les en-têtes de colonne sont les suivants : :

- Nom : Indique le nom de l'instance.
- n : Indique la taille du problème.
- Φ^* : La meilleure solution connue.
- LB : La meilleure valeur (borne inférieure) obtenue par chaque stratégie.
- $Iter$: Nombre d'itérations pour lesquelles la meilleure valeur LB est obtenue (limitée à 500).
- $GAP = \dfrac{\Phi^* - LB}{\Phi^*}$.
- CPU : Temps total d'exécution, en seconde pour calculer la meilleure valeur LB obtenue par chaque stratégie.

Nom	n	Ø*	Pure				MGT				ADS				NMDS			
			LB	Gap	Iter	CPU(s)	LB	Gap	iter	CPU(s)	LB	Gap	Iter	CPU(s)	LB	Gap	Iter	CPU(s)
tsp5	5	148	148	0	5	0.089414	148	0	5	0.035440	148	0	10	0.011177	148	0	5	0.014877
tsp6	6	207	207	0	5	0.052775	207	0	3	0.067979	207	0	3	0.015111	207	0	3	0.014534
tsp7	7	106.4	106.4	0	4	0.093246	106.4	0	4	0.040431	106.4	0	10	0.017036	106.4	0	4	0.015757
tsp10	10	378	378	0	20	0.139573	378	0	20	0.097444	378	0	40	0.056660	378	0	20	0.034543
Burma14	14	30.8786	30.8786	0	36	0.210967	30.8786	0	21	0.127149	30.8786	0	31	0.116396	30.8786	0	10	0.045628
tsp33	33	10861	10861	0	30	1.208825	10861	0	31	0.280733	10861	0	31	0.253657	10861	0	14	0.135275
eil51	51	426	400.1463	0.0607	100	6.552345	415.4136	0.0249	100	1.882163	412.1671	0.0325	100	1.836927	420.7088	0.0124	100	1.749168
st70	70	675	635.2332	0.0589	300	28.644063	619.8302	0.0817	300	10.546306	624.7722	0.0744	300	10.069613	635.2950	0.0588	300	10.125181
eil76	76	538	538	0	60	13.443316	538	0	20	0.768902	538	0	20	0.731305	538	0	18	0.724239
eil101	101	629	629	0	46	26.309066	629	0	50	3.344198	629	0	50	3.392173	629	0	50	3.380537

FIGURE 6. Résultats numériques

RÉFÉRENCES

[M.S. Bazaraa (1981)] M.S Bazaraa, H.D., Sherali, *On the choice of step sizes in subgradient optimization. Europ. J. Oper. Res. 7, 380-388, 1981.*

[M.S. Bazaraa (2006)] M.S. Bazaraa, H.D. Sherali, C.M. Shetty, *Non linear programming : Theory and algorithms. Wiley-Interscience series in discrete mathematics and optimization, 2006.*

[R. Belgacem (2017)] R. Belgacem, A. Amir, *A new modified deflected subgradient method. Journal of King Saud University - Science, In Press, 2017.*

[P.M. Camerini(1975)] P.M. Camerini, L. Fratta et F. Maffioli, *On improving relaxation methods by modified gradient techniques. In : Balinski M.L., Wolfe P. (eds) Nondifferentiable Optimization. Math. Program. Studies. 3. Springer, Berlin, Heidelberg., 1975*

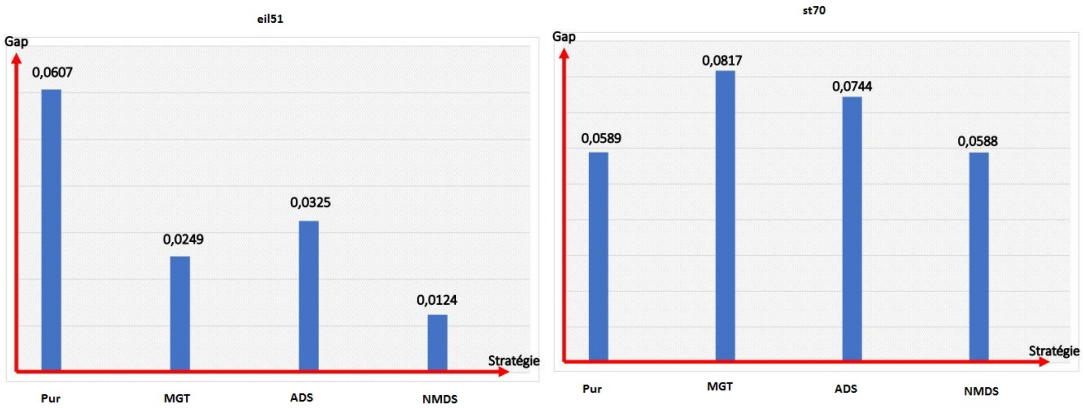


FIGURE 7. Représentation du Gap

[R. Fletcher (1964)] R. Fletcher, C.M. Reeves, *minimization by conjugate gradients. Comput. J. 7, 149-154, 1964.*

[M.H. Held(1974)] M.H. Held, P. Wolfe, H.D. Crowder, *Validation of subgradient optimization. Math. Program, 6, 62-88, 1974.*

[G. Laporte (1992)] G. Laporte, *The traveling salesman problem : an overview of exact and approximate algorithms. European Journal of Operational Research, 59, 231-247, 1992.*

[J. Nocedal (1999)] J. Nocedal and S.J. Wright, *Numerical Optimization. Springer- Verlag New York, Inc., 1999.*

[G. Reinelt (1994)] G. Reinelt, *The traveling salesman problem : computational solutions for TSP applications. Lecture Notes in Computer Science 840, Springer Verlag, Berlin, 1994.*

[A. Schrijver (1986)] A. Schrijver, *Theory of linear and Integer Programming, Wiley, 1986.*

[H.D. Sherali(1989)] H.D. Sherali, O. Ulular, *A primal-dual conjugate subgradient algorithm for specially structured linear and convex programming problems. Appl. Math. Optim. 20, 193-221,1989.*

Semantic-Similarity based Outlier Detection in Textual Data

Farek Lazhar
Computer Science Department
University 08 Mai 1945 - Guelma
Algeria
farek.lazhar@univ-guelma.dz

***Abstract*—In text-mining area, outlier documents have a negative impact on the performance of classification techniques, and detecting them remains a very challenging problem due especially to the curse of sparsity and high-dimensionality of textual data. In this paper, we propose an approach based on semantic similarity and that to detect outlier. Firstly, data is modeled using the Doc2Vec framework. Secondly, using Word Mover's Distance (WMD) a semantic similarity matrix is built. Thirdly, a statistical method is applied to identify outlier documents based on the sum of distances with all other documents. To show the effectiveness of our approach, two classification tests, one with original datasets and the second without outlier are applied. Experimental results show that discarding outlier from datasets conducts to improve the performance of classifiers.**

Keywords—Outlier Detection, Text Data, Doc2Vec Modeling, Sparsity, High-dimensionality, Classification

I. INTRODUCTION

Outlier detection, also called anomaly detection is the process that identify divergent observations i.e. observations which are not strongly related to the majority of observations in the same dataset. Outlier is defined as the set of objects that are considerably dissimilar from the remainder of the data [1]. Outlier is generally a data point which is different from the normal behavior of data points [2]. Also defined as a data value that seems to be out of place with respect to the rest of data [3]. Due to its effectiveness in data-mining area, outlier detection has been widely studied and attracted much attention of researchers in several domains including: defense, fraud detection, agriculture, etc. Various researches have been conducted on the outlier detection and their application in various domains [2-4].

Textual data is often characterized by its high dimensionality where redundant and irrelevant features are often present. To deal with the problem of sparsity and high dimensionality, topic modeling techniques are used such as Latent Drichlet Allocation (LDA), Latent Semantic Analysis, Non-Negative Matrix Factorization (NNMF), and recently Doc2Vec which proved its effectiveness in capturing better semantic similarity between documents.

In text mining, outlier documents carry much noise and make distance far from discriminative documents, hence outlier documents are considered as misleading for a classifier because of the high level of ambiguity they carry, which finally decrease its performance. However, classical distance metrics proved their inefficiency for measuring proximity in high dimensionality by showing surprising behavior [5].

To show the effectiveness of measuring semantic similarity in textual data, in this paper, we propose to combine Doc2Vec framework and Word Mover's Distance (WMD) as a distance metric to identify outlier documents. The remainder of this paper is organized as follows. Section 2 presents a literature review on some related works. In Section 3, the proposed approach is presented. Experimental datasets are described in section 4. Empirical results are discussed in Section 5. Our research work is concluded in Section 6.

II. LITERATURE REVIEW

Outlier Detection is widely studied and attracted the attention of several researchers, and due to the vagueness of outlier techniques, in this section, we focus on some similar methods that use distance metrics to identify outlier.

A distance-based algorithm [6] based on large dataset partition has been designed for outlier detection. Deciding either a point is outlier or not is based on computing the distance from its k^{th} nearest neighbor and then top n points are considered as outliers. Firstly, candidate partitions are generated using BIRCH's pre-clustering algorithm [7]. Secondly, an algorithm called nested-loop algorithm is used for computing outliers from the candidate partitions. Empirical results show its efficiency in outlier detection with respect to both data set size and data set dimensionality outperforming the nested-loop algorithm and another one called index-based algorithm.

In the same context, a distance-based approach [8] has been proposed for detecting outlier in large high-dimensional datasets. Similar to the proposed algorithm [6], a designed algorithm called HilOut is designed to efficiently detect the top n outliers where the sum of the distances separating a data point from its k nearest-neighbors is used as a weighting scheme. Data is linearized using the notion of space filling curve. Two functions called temporal cost and special cost respectively are used in the first phase to provide an approximate solution. In this phase, the algorithm iteratively isolates candidate data points to be outliers and at once reduce the dataset size. The algorithm stops when the dataset size reaches n. In the second phrase, an

exact solution is provided examining further the candidate outliers that remained after the first phase. Tested on large high dimensional datasets, the proposed algorithm always reports good solutions after a finite number of iterations.

A distance-based algorithm called ODDC (Distribution Clustering Outlier Detection) [9] has been proposed for outlier detection. It is based on mapping data into a new feature space, where the transformation vector captures the distance distribution of each point. ODDC proved its efficiency compared to a well-known outlier detection algorithm called LOF (Identifying Density-Based Local Outliers) [10] regardless the size of datasets, the dimensionality and the percentage of outliers.

Some known algorithms that use embedded distance metrics such as K-Means and its variants: K-Medoids, K-Medians, K Modes, Fuzzy K-Means [11], etc., are widely used in the literature of outlier detection. Numerous other variants of these algorithms such as BIRCH (Balanced Iterative Reducing and Clustering Hierarchies) [7], PAM (Partitioning Around Medoids) [12], CLARA (Clustering LARge Applications) [13] and CLARANS (Clustering Large Applications based upon RANdomized Search) [13] are also used. PAM clustering algorithm [12] has been applied considering that small-sized clusters are good holders for outlier objects. An algorithm called I-CLARANS [12] has been proposed, which is indeed a modified variant of CLARANS algorithm [13] using some geometric proprieties, and that to identify outlier. Empirical results show that I-CLARANS algorithm performs better in detecting outlier compared to PAM, CLARA and CLARANS.

III. Proposed Approach

In this work, brut textual data is mapped from a sparse and high-dimensional space to a compact vector one using the Doc2Vec framework, and that to capture semantic similarities between representative vectors when applying WDM metric. This transformation of data can be seen as an important process which decreased the difficulty of working with textual data compared to other approaches that work only with quantitative normally (Gaussian) distributed data.

A. Problem Overview

Given a corpus of *n* documents $D = \{d_1, d_2, \dots, d_n\}$, where each document d_i is assigned to one target class from $C = \{c_1, c_2, \dots, c_k\}$, *k* is the number of classes.

Let *cl* a classification technique. The goal of this work is to detect a subset *S* from *D* that poorly affect the performance of *cl*. By removing *S*, the accuracy of *cl* should be performed compared to its performance before removing *S*. i.e., discarding outlier documents from *D* can help to improve the accuracy of *cl*.

A distance score $\hat{s}$ is computed for each document *d* based on the sum of distances with all other documents, if $\hat{s}$ deviates obviously from the rest of scores then *d* is considered as outlier.

Figure 1 shows the main steps of our approach.

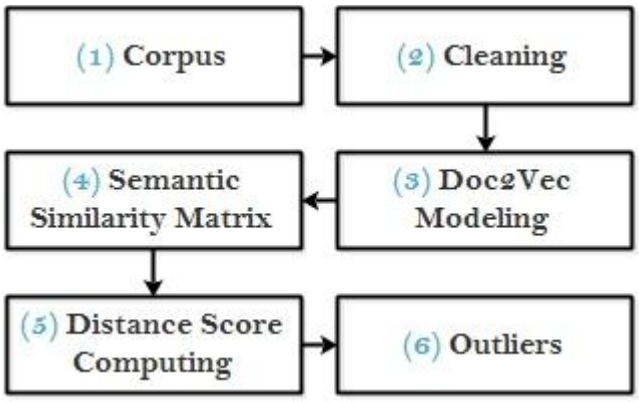


Fig. 1. General overview of the proposed approach.

B. Data Cleaning

Text Data may carry much irrelevant and undesirable contents such as stop-words, punctuation, HTML tags, etc., hence, and for an effective content, cleaning text data is very recommended to reduce the sensitivity of semantic similarity metrics to the high-dimensionality.

Stop-words are words that are frequent and do not carry any meaning, including articles, prepositions and some other high frequency words, such as 'a', 'the', 'of', 'and', 'and', 'it', 'I', 'you', 'that', 'this', 'those', etc. Punctuation and special chars are also high-frequent and do not carry any meaning such as ?, !, %, $, >, #, &, etc.

HTML tags are keywords surrounded by angle brackets like <html>, <head>, , , etc., which do not carry semantic information. By removing those frequent tags, the dimension of the corpus should be considerably reduced.

C. Topic modeling with Doc2Vec

Text Data is characterized by its high dimensionality and sparsity which poorly affect the measuring of semantic similarity between documents. Distance measures like the Euclidean distance for high dimensional data exhibit surprising properties that differ from what is usual for low-dimensional data [15]. To overcome the problem of sparsity and dimensionality, semantic modeling methods are used.

Doc2Vec, also called Paragraph Vector is one of the most effective data modeling technique to learn document-level embedding and represent documents as a vector, introduced in 2014 by Le and Mikolov [10], which is in fact a generalizing of Word2Vec method [16], and that by extending the learning of embedding from words to word sequences. Word2Vec is a three-layer neural net with one input, one hidden and an output layer. It implements CBOW (Continuous Bag of Words) and SkipGram architectures for computing vector representations of words, including their context [17]. CBOW tries to predict a word on bases of its neighbors i.e. predicting a word given its context (Syntactic Relation). However, SkipGram tries to predict the neighbors of a word i.e. predicting the context given a word (Semantic Relation).

Formally, it described as follows, every word is mapped to a unique vector, represented by a column in a matrix *W*. Given a sequence of training words $w_1, w_2, \ldots, w_T$, the objective of the word vector model is to maximize the average *log* probability [13]. Given a word *w* and its surrounding (context) words, CBOW and SkipGram respectively optimize the following objective functions :

$$\zeta_{CBOW} = \sum_{w \in W} \log\, p\big(w \,/\, Context(w)\big) \tag{1}$$

$$\zeta_{SkipGram} = \sum_{w \in W} \log\, p\big(Context(w) \,/\, w\big) \tag{2}$$

Doc2Vec explores Word2Doc framework by adding additional input nodes representing documents as additional context. Each additional node can be thought of just as an identifier for each input document. The objective of Doc2Vec learning is:

$$\max \sum \log p\big(tar \,/\, (con, doc)\big) \tag{3}$$

Where: tar: target word, con: context words, doc: document context.

In our study, choosing Doc2Vec for topic modeling is motivated by the empirical evaluation [18] which shows its effectiveness when it is trained on large corpora compared to other embedding methodologies. Also, the comparative study of semantic modeling methods [17] shows that Doc2Vec outperformed other semantic modeling methods such as LSA and LDA

D. Semantic Similarity Matrix (SSM)

After molding text data using Doc2Vec framework where documents are represented within normalized vectors with equal length. Now, it is possible to measure similarity between representative vectors using a distance metric such as Cosine distance, Euclidian distance, Manhattan distance, etc. However, in most of high dimensional applications, the choice of the distance metric is not obvious; and the notion for the calculation of similarity is very heuristical [5].

In this work, a distance measure called Word Mover's Distance (WMD) [19] which proved its efficiency in capturing semantic relations between text documents will be used to build our Semantic Similarity Matrix.

As described in [19], the distance between two text documents *A* and *B* is the minimum cumulative distance that words from document *A* need to travel to match exactly the point cloud of document *B*.

Considering the embedded matrix $X \in \mathbb{R}^{d \times n}$, provided with Word2Doc [16], for a finite size vocabulary of *n* words. The i^{th} column $x_i \in \mathbb{R}^d$, represents the embedding of the i^{th} word in d-dimensional space, assuming that text documents are represented as normalized bag-of-words (nBOW) vectors, $d \in \mathbb{R}^n$. The distance between words *i* and *j* are the Euclidean distance of their embedded word vectors x_i and x_j respectively, denoted by:

$$d(i,j) = \left\| x_i - x_j \right\| \tag{4}$$

The document distance, which is WMD here, is defined by:

$$\sum_{i,j} T_{ij}\, c(i,j) \tag{5}$$

Where T is a n × n matrix. Each element $T_{ij} \geq 0$ denotes how much of word *i* in the first document (denoted by *d*) travels to word *j* in the new document (denoted by d'). Then the problem becomes the minimization of the document distance, formulated as:

$$min_{T \geq 0} \sum_{i,j=1} T_{ij}\, c(i,j) \tag{6}$$

given the constraints:

$$\sum_{j=1}^{n} T_{ij} = d_i \tag{7}$$

$$\sum_{i=1}^{n} T_{ij} = d'_j \tag{8}$$

Now, and in order to construct the SSM, for each pair of documents (i, j) , we note by $w(i,j)$ the computed WMD between the document *i* and the document *j*. The constructed matrix is illustrated in Table 1.

TABLE I. SEMANTIC SIMILARITY MATRIX

	d_1	d_2	…	d_n
d_1	1	w_{12}		w_{1n}
d_2	w_{21}	1		w_{2n}
. . .				
d_n	w_{n1}	w_{n2}		1

The constructed symmetrical and square matrix has the following characteristics:

- Each diagonal entry equal to 1, which means that the semantic similarity between a document and itself is equal to 1 (100%).
- Each off-diagonal entry is between 0 and 1, which means that the similarity varies between 0 (completely non identical) and 1 (completely identical).

E. Similarity Score Computing

Using SSM computed in the previous subsection, for each document d_i from our data set *D*, a similarity score $\hat{s}$ is computed using the following formula:

$$\hat{s} = \frac{\sum_{j=1}^{N} w(d_i, d_j)}{N} \tag{9}$$

Where N is the total number of documents, $j \in [1, N]$, and $w(d_i, d_j)$ is the WMD between d_i and d_j.

F. Outlier Detection

Now, since each document is represented by one single real value which is the distance score computed in the previous subsection, we can consider that a document where its score deviates from the rest of scores is considered as an outlier.

A common statistical method using the standard deviation is applied to identify distant points (documents vectors) that are further away from the mean. Let m and δ the mean and the standard deviation of scores respectively. For N candidate data points, m and δ are given by the following formulas:

$$m = \frac{1}{N}\sum_{i=1}^{N} \hat{s}(i) \tag{10}$$

$$\delta = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{s}(i) - m)^2} \tag{11}$$

Standard deviation-based outlier detection method removes points that are above ($m + 2 \times \delta$) and points that are below ($m - 2 \times \delta$) .

IV. EXPERIMENTAL DATASETS

Three datasets [21] have been used to evaluate the performance of our approach. Each dataset contains 1000 sentences labelled with positive or negative sentiment, extracted from reviews of products, movies, and restaurants. The sentences come from three different websites: *imdb.com*, *amazon.com* and *yelp.com*.

As shown in Table 2, for each dataset the number of reviews is equally distributed over positive and negative classes.

TABLE II. EXPRIMENTAL DATASETS

Dataset	***#Reviews***	***#Positives***	***#Negatives***
Yelp	1000	500	500
Amazon	1000	500	500
IMDB	1000	500	500

V. EXPERIMENTS AND RESULTS

Using datasets mentioned in section 4. The first phase is dedicated to test the ability of our approach in mining outlier documents where data cleaning and Doc2Vec modeling are accomplished, Semantic Similarity Matrix is constructed and Distance Scores are computed. In the second phase, classifiers are tested. At first, with original datasets (i.e. with outlier documents) and the performance in term of F1-measure is computed. Then, the same classifiers without outlier documents are applied to show the effect of outlier documents on the behavior of classification algorithms.

In the second phase, three popular classification techniques commonly used in text classification are used: Random Forest (RF), Naive Bayes (NB) and Stochastic Gradient Descent Classifier (SGD). A detailed description of these three techniques can be found in [20]. We evaluated the performance of chosen classifiers with the F1-score which is given by the following formula:

$$F1 = 2 \times \frac{p \times r}{p + r} \tag{12}$$

Where, Precision $p = tp/(tp + fp)$ is the fraction of all positive predictions that are actual positives. Recall $r = tp/(tp + fn)$ is the fraction of all actual positives that are predicted to be positive. True positives *tp*, false positives *fp*, false negatives *fn*, and true negatives *tn* are represented via a confusion matrix (Table 3).

TABLE III. CONFUSION MATRIX

	Actual Positive	***Actual Negative***
Predicted Positive	*tp*	*fp*
Predicted Negative	*fn*	*tn*

Outlier documents are identified after mapping data into a numerical space by associating each document by its Similarity Score. Representative scores that obviously deviate from the mean are considered as outliers. Table 4 shows the results of classification before and after outlier removal.

TABLE IV. CLASSIFICATION RESULTS BEFORE AND AFTER OUTLIER REMOVAL

Dataset	***Classifier***	***F1***	
		Before	***After***
Yelp	RF	0.7410	0.7614
	NB	0.8032	0.8173
	SGD	0.6881	0.6896
Amazon	RF	0.7980	0.7983
	NB	0.7416	0.7813
	SGD	0.7874	0.7960
IMDB	RF	0.7316	0.7907
	NB	0.8078	0.8158
	SGD	0.7359	0.7781

Table 5 shows for each dataset the total number of reviews (sentences) and the number of detected outliers for both positive and negative classes. As shown in Table 4, despite the slight improvement recorded after outlier removal, we can say that outlier documents have a negative influence on classification algorithms when comparing F1-measure before outlier removal.

TABLE V. OUTLIER DETECTION RESULTS

Dataset	*#Reviews*	*#Outliers*		
		Positive	*Negative*	*Total*
Yelp	1000	20	16	**36**
Amazon	1000	22	25	**47**
IMDB	1000	21	25	**46**

Although our approach is time-consuming when the number of documents is important, the challenging problems of high dimensionality and sparsity have been overcome using Doc2Vec topic modeling framework and that to make useful the use of distance metrics when calculating similarity between representative Doc2Vec vectors.

For normal distributed data, several transformation techniques can be applied. In [8], dataset is transformed using space-filling curve as a linearization technique which make sense to the distance between data points which then helped to detect outlier points. Very similar to [8], the distance based approach [9] show that for each data point computing distances distributed over clusters can help to detect outliers. This method transits the feature space to the new space by discretizing the distance distribution of each object. Where each point in the original space is represented as a new vector, and each dimension of the new vector is a distance distribution. After clustering, objects in small clusters as considered as outliers.

However, these methods of transformation cannot be applied to textual data, where distance measurements proved the inefficiency when measuring semantic similarity between documents. The problem has been resolved by using the Doc2Vec framework which eliminates the issue of sparsity and high-dimensionality, therefore capturing semantic similarity became possible.

VI. CONCLUSION

In this paper, we proposed a combined approach based on the Doc2Vec topic modeling framework and Word Mover's Distance (WMD) in order to detect outlier in textual data. We faced two main challenging problems in the text mining area which are high-dimensionality and sparsity. Using Doc2Vec to represent data into a semantic vector space and reduce dimensionality, the distance calculation between representative vectors became useful.

For each document, a semantic similarity score is computed. Then, a document represented by its score is considered as an outlier when it deviates considerably from the mean of all other scores.

REFERENCES

[1] J. Han and M. Kamber, "Data Mining: Concepts and Techniques", 743. Morgan Kaufmann,San Francisco, 2006.

[2] J. Tamboli and M. Shukla, "A survey of outlier detection algorithms for data streams", 3rd International Conference on Computing for Sustainable Global Development (INDIACom), 2016, pp 3535 – 3540.

[3] SS. Sreevidya, "A Survey on Outlier Detection Methods", International Journal of Computer Science and Information Technologies (IJCSIT), vol. 5 (6), pp 8153 – 8156, 2014.

[4] S. Sharma and R. Jain, "Outlier Detection in Agriculture Domain: Application and Techniques". In: Aggarwal V., Bhatnagar V., Mishra D. (eds) Big Data Analytics. Advances in Intelligent Systems and Computing, vol. 654. Springer, Singapor, 2018.

[5] A. Blum and T. Mitchell, "Combining labeled and unlabeled data with cotraining". In: Proceeding COLT' 98 Proceedings of the eleventh annual conference on Computational learning theory, pp 92 – 100, 1998.

[6] E.M Knorr and R.T Ng, "Algorithms for mining distance-based outliers in large datasets". In: Proceeding VLDB Algorithms for mining distancebased outliers in large datasets, pp 392 – 403, 1998.

[7] G. Singh and V. Kumar, "An Efficient Clustering and Distance Based Approach for Outlier Detection", International Journal of Computer Trends and Technology (IJCTT), vol. 4 (7), pp 2067 – 2072, 2013.

[8] F. Angiulli and C. Pizzuti, "Outlier Mining in Large High-Dimensional Data Sets", IEEE Transactions reon Knowledge and Data Engineering, vol. 17 (2), pp 203–215, 2005.

[9] K. C. Niu, S. Huang, S. Zhang, and J. Chen, "ODDC: Outlier Detection Using Distance Distribution Clustering", T. Washio et al. (Eds.): PAKDD 2007 Workshops, Lecture Notes in Artificial Intelligence (LNAI) 4819, Springer-Verlag, pp 332 – 343, 2007.

[10] M.M. Breunig, H. Kriegel, R.T. Ng, et al., "LOF: Identifying Density Based Local Outliers". In: Proceedings of ACM SIGMOD International Conference on Management of Data, Dalles, TX, pp 93 –104, 2000.

[11] J.C. Bezdek, "Pattern recognition with fuzzy objective function algorithms", New York: Plenum Press, 1981.

[12] V. Kumar, S. Kumar and A.K Singh, "Outlier Detection: A Clustering Based Approach", International Journal of Science and Modern Engineering (IJISME), vol. 1 (7), pp 16 – 19, 2013.

[13] Q. Le and T. Mikolov, "Distributed Representations of Sentences and Documents", ICML'14 Proceedings of the 31st International Conference on International Conference on Machine Learning, vol. 32, pp 1188 – 1196, Beijing, China, 2014.

[14] G. Singh and V. Kumar, "An Efficient Clustering and Distance Based Approach for Outlier Detection", International Journal of Computer Trends and Technology (IJCTT), vol. 4 (7), pp 2067 – 2072, 2013.

[15] K. Klawonn, F. Hoppner, K. Shim and B. Jayaram, "Efficient algorithms for mining outliers from large data sets". Proceeding Revised Selected Papers of the First International Workshop on Clustering High Dimensional Data, vol. 7627, pp 14 – 33, 2013.

[16] T. Mikolov, K. Chen, G. Corrado and J. Dean, "Efficient estimation of word representations in vector space". In Proceedings of Workshop at the International Conference on Learning Representations, Scottsdale, USA, 2013.

[17] M. Campr and K. Jezek, "Comparing Semantic Models for Evaluating Automatic Document Summarization". In: Kral P., Matousek V. (eds) Text, Speech, and Dialogue. TSD 2015. Lecture Notes in Computer Science, vol. 9302. Springer, Cham, 2015.

[18] J.H. Lau and T. Baldwin, "An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation". Proceedings of the 1st Workshop on Representation Learning for NLP, Berlin, Germany. pp 78 – 86, 2015.

[19] M. Kusner, Y. Sun and N. Kolkin, "Weinberger K. From word embeddings to document distances". In: International Conference on Machine Learning, pp 957 – 966, 2015.

[20] D. Forsyth, "Learning to Classify", In: Probability and Statistics for Computer Science, Springer, Cham, 2018.

[21] K. Dimitrios, D. Misha, D.F Nanado and S. Padhraic, "From Group to Individual Labels using Deep Features", In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, NSW, Australia, pp 597 – 606 ,2015.

A comparative Study Between Merkle-Damgard And Other Alternative Hashes Construction

Amine Zellagui
dept of electronic,University of Sciences and Technology of Oran
laboratory of coding and security of information LACOSI
Oran, Algeria
Email : amineget29@gmail.com

Naima Hadj-Said
dept of computer science,University of Sciences and Technology of Oran,
laboratory of coding and security of information LACOSI
Oran, Algeria
Email : nim_hadj@yahoo.fr

Adda Ali-Pacha
dept of electronic,University of Sciences and Technology of Oran
laboratory of coding and security of information LACOSI
Oran, Algeria
Email : a.alipacha@gmail.com

***Abstract*— Hash functions are used to verify the integrity and authenticity of information and play fundamental role in cryptography and web application, The most traditional hash function such as MD4, MD5, SHA-1, SHA-2 is based on Merkle-Damgård construction which a compression function is used iteratively. this construction proves that the security of hash function relies on the security of the compression function, However ,certain generic attacks exist that differentiate an MD hash function such as the multi-collision attack and length extension attacks. in this paper ,We present a comparative study in terms : security requirements and different design methods of compression function between Merkle-Damgård construction and other alternative recent hash construction like MDP, Sponge, Wide-pipe and HAIFA, which represented by various hash functions specifications (Keccak ,BLAKE ,grostl ,JH..) in the NIST SHA-3 competition.**

Keywords*— *Cryptography ,Hash Function, Hash Construction ,MD ,SHA-3 ,NIST

I. INTRODUCTION

The Internet has become the main means of communication. Thus, the data of the user reaches the highest priority in the field of data communication. To maintain reliable network usage, data integrity, data authentication, non-repudiation, data privacy is of utmost importance. Hashing is an important technique used for secure communication in the presence of indiscretions. It provides all the essential aspects of information security such as integrity, authentication and confidentiality. The password hash is lightweight and convenient to use and can defend against phishing attacks [1]. A hash algorithm takes a piece of data and produces a hash: an irreversible string of a fixed length. Because the hash cannot be returned in the original message, it is used to check the data, rather than decode it. Thus, hashing is an essential part of password protection. If a system uses a hash algorithm for verification rather than processing actual data directly, security is improved because there is only a brief window of time when the original data is used.

Collision resistance is a property of cryptographic hash functions, a cryptographic hash function H is collision resistant if it is difficult to find two entries that give the same hash value

In recent years, many traditional hash functions based on Merkle Damgård construction such as MD4 [2] , MD5 [3] ,has been successfully attacked by X Wang in 2004 [4] and SHA-1 [5] by Marc Stevens in 2017 [6]. So, It encourages researchers to find alternatives to synthesis efficient hash functions with ease of computations.

II. HASH FUNCTIONS IN CRYPTOGRAPHY

Cryptographic systems are obtained by combining the use of primitives [8], among which there are often hash functions, as in (1).

The hash functions are used to calculate an arbitrary size input data with a fixed size fingerprint. This size generally varies between 128 and 512 bits.

$$. \quad H: \{0,1\}^* \rightarrow \{0,1\}^n \qquad (1)$$

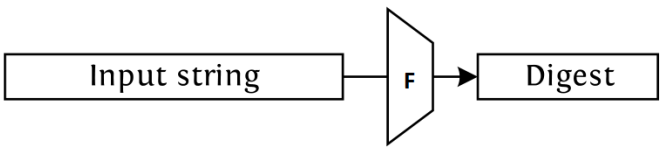


Fig.1Hash function.

Originally, hash functions were created to facilitate database management. Rather than manipulating data of variable and potentially large size, these data are associated with a fixed-size fingerprint, for which comparisons are quicker to make.

Properties Hash functions are not cryptographic systems : they alone do not provide a security property. In particular, they do not constitute an encryption algorithm [3].

A. No key

The definition of hash functions does not involve a key. For some applications, the hashed data is entirely public, and a prospective attacker can perform the hash calculations himself".

B. Absence of inversion algorithm

Unlike block-based or asymmetric encryption primitives, hash functions do not need to be inverted. The impossibility of calculating an antecedent by a hash function is even a fundamental security property of these primitives.

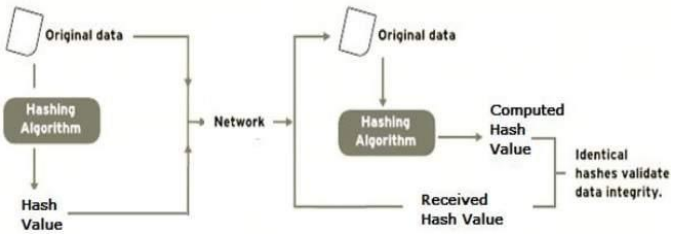


Fig. 2 Principle of operation of a hash function

A hash function have good cryptographic properties if it is resistant to preimage, second-preimage and collisions. The following three problems must therefore be difficult.

- *Preimage :* Given a randomly chosen hash, find a message m such that H (m) = h.
- *Second preimage :* Given a message m chosen randomly, find a message m' such that H (m) = H (m').
- *Collision :* Find two messages m, m' , such that m ≠ m' and H (m) = H (m ').

III. MERKLE-DAMGARD CONSTRUCTION

The most traditional hash function like MD4, MD5, SHA-1, SHA-2 [8] use the Merkle Damgård construction. its designed by Merke [9] and Damgård [10] in 1989. a compression function is used iteratively. The message M, after having been prepared, is divided into blocks of fixed size m1, ..., mk and the compression function h processes the blocks one after the other.

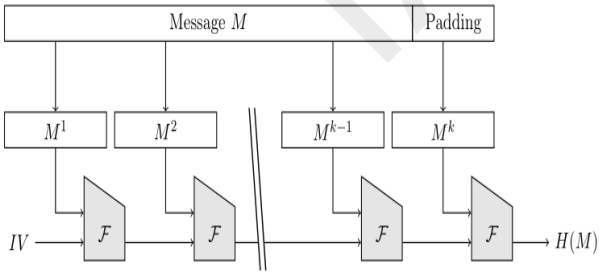


Fig. 3 Merkle-Damgård construction

where :

Padding is the bytes added to the message M ,F is the compression function, IV is initial vector and H is final hash.

The Merkle-Damgård construction proves that the security of hash function relies on the security of the compression function, However ,certain generic attacks exist that differentiate an MD hash function such as the multi-collision attack [11] and length extension attacks.

Attack by length extension : This attack, which is one of the main weaknesses of the Merkle-Damgàrd building, makes it possible to build from the imprint H (m) of a message m, the hash of a message m' which is suffix of m , without knowing the message itself.

Let m and m' be two messages, such as pad (m) or a pad prefix (m '). There are then indices k, ℓ such that :

$$\text{pad}(m) = m1 \,||m2\, ||...||\, m_k$$
$$\text{pad}(m') = m1||m2||...||m_{k+1}\, ||...||\, m\ell$$

We can now construct the hash of m' as a function of h_k= H (m) by calculating the following:

$$h_i = h\,(\, h_{i-1}\,, m_i\,) \text{ avec } i \geq k + 1.$$

This attacks has encouraged researchers to use other hash construction

Attack by Multi collisions : In an iterated hash function, multi-collisions can be constructed more efficiently using internal collisions. We start by looking for a pair of messages m_0, m_0' which gives an internal collision starting from the IV. Then we look for a pair of messages m1, m1 'which gives an internal collision starting from the state x0 after having processed m0. If we repeat this operation k times, we get k internal collisions, and we can build 2^k messages by choosing one or the other messages at each step

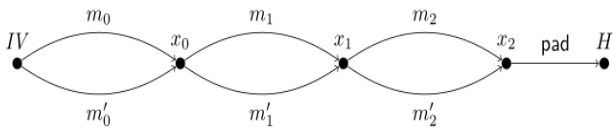


Fig. 4 Multi collisions attack

IV. ALTERNATIVES HASH CONSTRUCTION

1. Review of hash construction :

In this subsection we present some recent hash construction scheme.

A. Merkle-Damgård with permutation (MDP)

This construction is the modified version of Merkle-Damgård construction by Hirose [12].the idea is to adding a permutation before the last block.

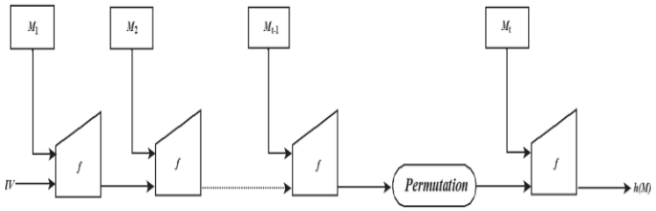


Fig. 5 Merkle-Damgård with permutation.

The researcher proved that the collision resistance of this construction follows trivially from the collision resistance of the Merkle-Damgård construction as the former introduces minimal changes to the latter, and is neither preimage nor second preimage resistant.

B. Wide pipe construction

The wide pipe constructions designed by Stephan lucks in [13] to avoid the length of extension attack, multi-collisions attack and other vulnerabilities of Merkle Damgård construction.

the idea is to increase the internal state L of the hash function and generate the output of length n-bits using second compression function F', where L > n

by increasing the size of the internal state, finding collisions for the compression function becomes harder, which complicates the other generic attacks.

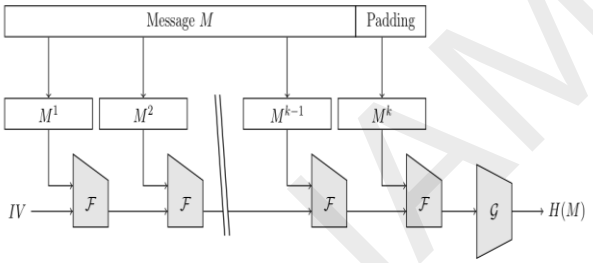


Fig. 6 Wide pipe construction

-Wide-pipe construction withstands extended length attacks and searches for multi collision.
There are several instances of wide pipe construction. Among the best known are the double pipe mode [13], defined by b = 2n, and the Chop-MD mode [14], where the G function is a simple truncation. Functions with b = n are also known as narrow pipe [15].

C. Tree Construction

The tree construction is designed by Ralph Merkle to support multi core platforms [16]. where, the message M is divided into block of fixed size m1, m2, m3,,mn. the compression function F take two block as input and generate a fixed size as output, every hash are then concatenated two by two to calculate a new hash and so on , Until getting the final hash value h.

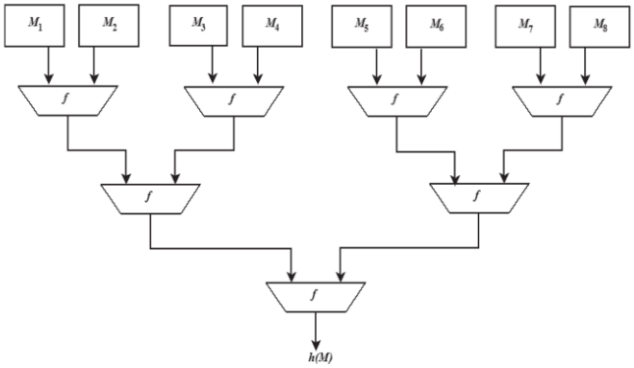


Fig. 7 Tree construction

D. Sponge construction

Another recently introduced construction called sponge construction is due to Bertoni [17]. It is an iterated construct capable of producing arbitrary size outputs. Unlike the Merkle-Damgàrd construction which is based on a compression function, the sponge construction relies on a fixed-size transformation, that is to say on a permutation f, operating on the b-bit words. In practice, all known instances of this construct are based on a permutation.

The procedure takes place in two successive stages.

The absorbing step : In this step, the blocks of the padded message are "absorbed" iteratively. The first block m1 is combined using an exclusive OR (XOR) with the state. The transformation f is then applied to the result of this operation. Then, the second message block m2 is added to the state and the transformation f is called again. The same procedure is repeated until all message blocks are absorbed.

The squeezing step : During this step, blocks of the internal state are extracted at separate locations by applications of the transformation f. The size of the blocks extracted, can be chosen by the user.

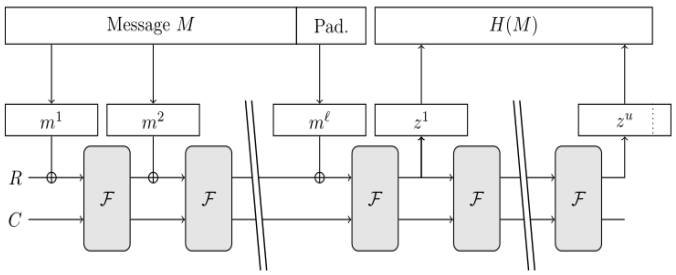


Fig. 8 Sponge construction

where :
r is the bit-rate and c is the capacity, Both are initially set to zero.
The security of the hash function can be reduced to the security of the compression function, but that does not mean that an attack on the compression function gives an attack on the hash function. In particular, if the internal state is sufficiently large compared to the output size n, it is possible to have a hash function that is resistant to collision or pre-image attacks even if it is easy to find collisions or pre-images when we control the chaining variable.
The sponge construct has been used as a domain extension algorithm for several newly introduced hash functions. Among them is the new SHA-3 standard.

E. HAIFA construction

The hash iterative framework (HAIFA) is designed by Biham in 2007 [18], to avoid the length extension attack. it's one of the alternative to Merkle Damgård construction. HAIFA maintains the good properties of the Merkle-Damgàrd construction while adding to the security of the transformation, as well as to the scalability of the transformation.

The idea is introducing extra input parameters to the compression function , which are a fixed salt value of s-bits along with a counter IV of t-bits to every message block in the iteration of the hash function.
The padding scheme used in HAIFA is very similar to the one used in the Merkle-Damgard construction: In HAIFA the message is padded with 1, as many needed 0's, the length of the message encoded in a fixed number of bits, and the digest size:

- Pad a single bit of 1.
- Pad as many 0 bits as needed such that the length of the padded message (with the 1 bit and the 0's) is congruent modulo n to (n − (t + r)).
- Pad the message length encoded in t bits.
- Pad the digest size encoded in r bits

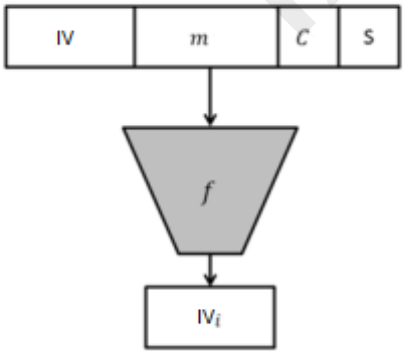


Fig. 9 HAIFA construction

this construction is used by BLAKE hash function which is one of the candidate in the NIST hash function competition

2. Security analysis :

Table 1 Security analysis between Merkle-Damgard and alternatives construction

	MD	MDP	Sponge	HAIFA	Tree	Wide-pipe
Length-extension	**yes**	**yes**	**no**	**no**	**no**	**no**
Internal collision	$2^{n/2}$	$2^{n/2}$	2^n	$2^{n/2}$	2^n	2^n
Multi collision	$2^{n/2}$	$2^{n/2}$	2^n	$2^{n/2}$	2^n	2^n
Second preimage	2^{n-}	2^{n-k}	2^n	2^n	2^n	2^n

from table 1,we can see that the sponge construction and wide-pipe have high complexity 2^n.

V. CONCLUSION

In this paper, a comparative study between Merkle-Damgård other recent construction like sponge, wide-pipe and HAIFA, which represented by various hash functions specifications in the NIST SHA-3 competition has been presented.
The security of hash function depends on the collision resistance property of the underlying compression functions. Merkle-Damgård construction method failed to preserve this important security property..
Sponge construction and keccak are the winners of NIST hash function competition, We recommended to use Sponge construction to create other hash function.

REFERENCES

[1] C.G Thomas, Robin Thomas Jose, A Comparative Study on Different Hashing Algorithms, Vol. 3, Special Issue 7, October 2015.

[2] Rivest R., 1992, "The MD4 Message-Digest Algorithm,"RFC 1320, MIT LCS and RSA Data Securit y,

[3] Rivest R., 1992, "The MD5 Message-Digest Algorithm,"RFC 1321, MIT LCS and RSA Data Securit y,

[4] Wang, X., Feng, D., Lai, X., & Yu, H. (2004). Collisions for hash functions MD4, MD5, HAVAL-128 and RIPEMD.Cryptology ePrint Archive, report 2004/199

[5] D. Eastlake ,P. Jones,"US Secure Hash Algorithm 1 (SHA1)",RFC3174,September 2001..

[6] Marc Stevens, Elie Bursztein, Pierre Karpman, Ange Albertini, Yarik Markov,The first collision for full SHA-1,CWI Amsterdam,Google Research 2017.

[7] SONG, Dawn Xiaoding, WAGNER, David, et PERRIG, Adrian. Practical techniques for searches on encrypted data. In : Security and Privacy, 2000. S&P 2000. Proceedings. 2000 IEEE Symposium on. IEEE, 2000. p. 44-55

[8] D. Eastlake,T. Hansen,2006, US Secure Hash Algorithms (SHA and HMAC-SHA) ,rfc4634.

[9] MERKLE, Ralph C. A certified digital signature. In : Conference on the Theory and Application of Cryptology. Springer, New York, NY, 1989. p. 218-238.

[10] DAMGÅRD, Ivan Bjerre. A design principle for hash functions. In : Conference on the Theory and Application of Cryptology. Springer, New York, NY, 1989. p. 416-427.

[11] A. Joux, "Multicollisions in Iterated Hash Functions: Application to Cascaded Constructions," Crypto'04, 2004, LNCS, vol. 3152, pp. 306-316

[12] S. Hirose, J. H. Park and A. Yun, "A Simple Variant of the Merkle-Damgård Scheme with a Permutation," Asiacrypt'08, 2008, LNCS, vol. 4833, pp. 113-129

[13] S. Lucks, "A Failure-Friendly Design Principle for Hash Functions," Asiacrypt'05, 2005, LNCS, vol. 3788, pp. 474-494

[14] BRESSON, Emmanuel, CANTEAUT, Anne, CHEVALLIER-MAMES, Benoıt, et al. Shabal, a submission to NIST's cryptographic hash algorithm competition. Submission to NIST, 2008.

[15] GLIGOROSKI, Danilo. Narrow-pipe SHA-3 candidates differ significantly from ideal random functions defined over big domains. NIST mailing list, 2010, vol. 2010.

[16] PRENEEL, Bart. The state of hash functions and the NIST SHA-3 competition. In : International Conference on Information Security and Cryptology. Springer, Berlin, Heidelberg, 2008. p. 1-11.

[17] BERTONI, Guido, DAEMEN, Joan, PEETERS, Michael, et al. On the indifferentiability of the sponge construction. In : Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Berlin, Heidelberg, 2008. p. 181-197.

[18] BIHAM, Eli et DUNKELMAN, Orr. A Framework for Iterative Hash Functions---HAIFA. Computer Science Department, Technion, 2007.

Application of Convex Optimization Results of DE FINETTI's problem for Proportional Reinsurance

Cheraitia Zahra
department of Statistics
The National School of Statistics and Applied Economics
Tipaza, Algeria
Cheraitia-zahra@outlook.fr

Abstract

The convex functions appear abundantly in the engineering and allow to model many nonlinear phenomena (physics equations, signal, game theory and economics, statistics, etc.). They have remarkable specificities that allows actuaries to minimize financial risks to which some institutions are exposed, especially insurance companies. Therefore, the use of mathematical tools to manage the various risks is paramount.

The objective of this work is to find the optimal retention level for a proportional reinsurance treaty based on the results of the convex optimization developed in De Finetti's model. The latter makes it possible to determine the level of retention that achieves the expected profit by the insurer, while minimizing claims volatility.

Keywords: convex function, nonlinear optimization, proportional reinsurance.

I. INTRODUCTION

The study of convex functions has provided a framework in which a whole class of nonlinear functional analysis problems can be solved, coming from various fields such as mechanics, economics, partial differential equations or even analysis digital. Given the difficulty of approaching non-linear problems in a rather general way, their modeling plays a very important role which motivated the autonomous development of the theory.

In finance, optimization algorithms are essential to define models that minimize risks. For this purpose, convex functions are the main tool used in risk management. For example, De Finetti's work on nonlinear optimization has proved that the results of convex optimization allow to define a dynamic approach to programming in insurance when it comes to determine an optimal retention. Afterwards, both mathematicians F.Glineur and J-F.Walhin even extended De Finetti's results on other reinsurance treaties.

As a matter of fact, the insurer is responsible for guaranteeing the various risks to which its customers are exposed and for compensating the claims that have occurred. However, the risks covered can have a strong influence on the insurer's security and possibly on its outcome. As a result, the insurance company applies itself to risk management techniques, namely reinsurance, that allows the insurer to transfer part of its risk for a premium. The relationships between insurers and reinsurers are based on a community of interest, each party seeks to determine the commitment that meets its own objectives, by setting the share of the risk to be retained[1]. This part represents his own retention. The latter represents one of the most sensitive parameters in reinsurance, as transactions between the two organizations are common, a poor appreciation of the level of retention has a direct impact on their profitability. To this end, each party must precisely establish a reinsurance program that optimizes its commitment.

This paper aims to define an optimal proportional reinsurance structure based on De Finetti's results on convex optimization in order to define an efficient reinsurance structure.

II. THEORY OF CONVEX FUNCTIONS

Convex functions are one of the most basic types of functions. A function $f : R_n \rightarrow R$ is convex if its domain is a convex set and for all x, y in its domain, and all $\lambda \in [0, 1]$, we have :

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$$

Convex functions has a lot of applications. It is used for proving some inequalities in easy manner. Also it has various applications in operation research, Quadratic and Geometric programming problems.[2]

Convex functions are particularly easy to minimize (for example, any minimum of a convex function is a global minimum). For this reason, there is a very rich theory for solving convex optimization problems that has many practical applications (for example, circuit design, controller design, modelling, etc.)

[1] J. Blondeau C.Partrat.; "Reinsurance: a technical approach"; economica edition; France, 2003, pp 14.

[2] P. Constantin Niculescu, L.E Persson ; "Convex function and their application, a contemporary approach", Springer edition, New York, 2004, pp145.

III. DE FINETTI'S RESULTS ON NON-LINEAR PROBLEMS

Historically, actuaries have focused on optimal reinsurance in a two-dimensional space. Their purpose is to compare and improve their strength and performance. It is obvious that the insurer is always looking to maximize its profit and minimize the standard deviation of this profit.

Consider a portfolio with n independent risks $X_1, \ldots, X_n$ against a premium $P_1, \ldots, P_n$. We will protect each risk by a proportional reinsurance with a percentage of transfer (1 - θ). The ceded premium is loaded by a technical load based on the expected value; it is written as follows:

$$P(\text{surrendered}) = (1+ \xi i^{re})(1 - \theta) E(X)$$

De Finetti suggested choosing a divested quantity (1- θi) that minimizes the variance of the insurer's result and keeps its expectation value constant.
The result of the insurer is given by:

$$Z(\theta) = \sum_{i=1}^{n} (Pi - (1+ \xi i^{re})(1 - \theta_i) E(X_i) - \theta_i X_i)$$

Θ is a vector of proportionally conserved quantities.

$$\Theta = (\theta 1, \ldots, \theta 2).$$

The determination of this vector will be a problem of minimizing the volatility of claims;

Under the constraints

$$\begin{cases} \text{Min Var } Z(\theta) \\ E(Z(\theta)) = K \\ \theta i \geq 0, \ i=1,\ldots,n \\ \theta i \leq 1, \ i=1,\ldots,n \end{cases}$$

Where K is the expectation value of the result chosen by the insurer.[3]

To solve this problem, it is necessary to understand the variance of claims by policy.

$$\text{Var}(Z(\theta)) = \text{Var}\left[- \sum_{i=1}^{n} \theta i \text{Var}(Xi) \right] = \sum_{i=1}^{n} \theta i^2 \text{Var}(Xi)$$

The expectation value of the result is given by E(Z(θ))

$$= \sum_{i=1}^{n} (Pi - (1 + \xi i^{re})(1 - \theta i) E(Xi) - \theta i E(Xi)$$

$$= \sum_{i=1}^{n} Pi - \xi i^{re}(1 - \theta i) E(Xi) - \theta i E(Xi)$$

The problem to be solved is therefore;

$$\text{Min} \sum_{i=1}^{n} \theta i^2 \text{Var}(Xi)$$

Under the constraints:

$$\begin{cases} \sum_{i=1}^{n} \xi i^{re}(1 - \theta i) \, E(Xi) = -k + \sum_{i=1}^{n} Pi - \sum_{i=1}^{n} E(Xi) \\ \theta i \leq 1 \ , & i= 1,\ldots, n \\ \theta i \geq 0 \ , & i= 1,\ldots, n \\ -\theta i \leq 0 \ , & i= 1,\ldots, n \end{cases}$$

The solution of this modeling is:

$$\Theta i = \min \left[1, \max \left(0, 1 - \frac{\lambda \mu E(Xi)}{Var(Xi)}\right)\right]$$

λ is the constant of the Lagrange multiplier.

IV. OPTIMIZATION STEPS

A- The probability of ruin

In actuarial science and applied probability ruin theory (sometimes risk theory collective risk theory) uses mathematical models to describe an insurer's vulnerability to insolvency/ruin. In such models key quantities of interest are the probability of ruin, distribution of surplus immediately prior to ruin and deficit at time of ruin.

Since the amount X of the insurer's expenses during a certain period is subject to fluctuation, while the premium received at the beginning of the period is certain and fixed, it may happen that X exceeds the amount of the premium. In this case, it will be said that the insurer is ruined, the manager will talk about loss on the farm. Obviously, significant operating losses lead to the insolvency of the insurance company, which is more the term ruin.
The introduction of the ruin function makes it possible to evaluate the risk for the operating loss to exceed a given value. Mathematically, it looks like this:

$$P(X > R + P) \leq \alpha$$

- R: the amount of the reserves of the company
- P: the amount of premiums
- X: the overall burden of the claim
- α: A minimum level fixed a priori

The first goal of determining the probability of ruin is to logically evaluate the wealth of the company, that is to say the probability that the scenario translating a failure is realized, and to estimate the initial level of reserves to make this probability sufficiently low.

The insurer may consider using reinsurance to minimize the likelihood of bankruptcy. In the case of a quota share, the minimization takes place as well;
We consider a homogeneous group of n risks. The random variable is the expenditure from risk no. I and the total amount of expenses of the group for the current year and therefore

$$X = Y_1 + Y_2 + \ldots + Y_n$$

[3] F.Glineur y and J. F. Walhin; "Finetti's Retention Problem for Proportional Reinsurance Revisited", Secura, Belgium, 2006, pp2.

It is assumed that the pure premium is charged proportionally to the expected value. For simplicity, it will be assumed that the technical loading rate applied by the reinsurer to the pure reinsurance premium is also equal to λ, and that there is no commission payment by reinsurance.

By applying the following assumptions:

(H1): The random variables $Y1$, ... Yn are independent and identically distributed.

Given the above ratings, the cumulative annual amount of claims rated X is:

$$X= Y_1 + Y_2 + ... + Y_n$$

X follows the binomial law of parameters n and p, which is noted $X \sim$ Ɓ (n, p) from which one deduces the mathematical expectation value and the variance of the random variable X:

$$E(X)= n\,p \qquad \text{et} \qquad Var(X)= n\,p\,q$$

Let us introduce the additional hypothesis (H2) below that will allow us to easily carry out mathematical calculations;

(H2): It is assumed that the number n of insureds is large enough for the law binomial Ɓ (n, p) can be approximated by the normal distribution:

$$N\ (m, \sigma 2) \text{ with } (m = n\,p \text{ and } \sigma = \sqrt{n\,p\,q})$$

This approximation will be possible provided that the product npq is sufficiently large.

Approximately:

$$Xi \sim N\ (n\,p, n\,p\,q)$$

The random variable θX follows approximately the law N(θ E (X), V (X)).

$$P\ (\text{ruine}) = P\ \{\theta X > R + \theta \Pi(X)\}$$

Defining the reduced centered random variable U by:

$$U= \frac{\theta X - E(\theta X)}{\sigma(\theta(X))};$$

$$P\ (\text{ruined}) = P\ \{U > \frac{R+\theta(1+\lambda)E(X)-\theta E(X)}{\theta\sigma(X)}\}$$

$$P\ (\text{ruined}) = P\ \{U > \frac{R+\theta\lambda E(X)}{\theta\sigma(X)}$$

Or again; $P\ (\text{ruined}) = 1- F_0\ (\frac{R}{\theta\sigma(X)} + \lambda \frac{E(X)}{\sigma(X)})$

Moreover, we have seen previously that the profit before reinsurance is equal to:

$$B= \Pi(X)- X$$

Its mathematical expectation value and its variance have for expression:

$$E(B) = \lambda\, E(X) \qquad \sigma(X) = \sigma\ (B)$$

If the factor of safety before reinsurance is greater than or equal to the value t desired by the insurer (which is judged 4), that is:

$$T= \frac{R+E(B)}{\sigma(X)} = \frac{X+\lambda E(X)}{\sigma(X)} \geq t$$

So, reinsurance is not necessary. Conversely, if the previous inequality is not verified, the insurer will have to resort to reinsurance.

We are now in this case and it is assumed that the insurer wishes to best adjust the full conservation for a given risk.

The insurer's annual profit after reinsurance, denoted B (θ), is given by:

$$B(\theta) = \Pi(X) - \theta X - (1 + \mu)(1 - \theta)\ E(X)$$

Or the technical bonus can also be written

$$\Pi\ (X) = (1 + \lambda)E(X)$$

Given the above assumptions, we deduce the expectation value and the variance:

$$E\ (B(\theta))= (\lambda - \mu)E(X) + \mu\theta E(X)$$

And

$$Var\ (B(\theta)) = \theta^2 Var(X)$$

Or the technical bonus can also be written Π (X) = (1 +λ)E(X). Given the above assumptions, we deduce the expectation and the variance: $0 \leq \theta \leq 1$

$$T(\theta) = \frac{R+E(B(\theta)}{\sigma(B(\theta))} = \frac{R+(\lambda-\mu)E(X)+\mu\theta E(X)}{\theta\sigma(X)}$$

As we are in the context of proportional share reinsurance, the insurer commission = reinsurance commission, and we write:[4]

$$T(\theta) = \frac{R+\theta\lambda E(X)}{\theta\sigma(X)}$$

Retention is sought from the retention coefficient T(θ) :

$$T(\theta) \geq 4 \Rightarrow \frac{R+\theta\lambda E(X)}{\theta\sigma(X)} \geq 4$$

$$\Theta < \frac{R}{(4\sigma(X)-\lambda E(X))}$$

Therefore the probability of ruin will allow us to determine the insurer's retention interval. For this purpose, we had the idea of combining the probability of ruin theorem with convex optimization, and this to determine accurately the insurer's retention.

B- De Finetti's optimization on proportional reinsurance

After defining the retention interval from the probability of ruin, De Finetti's model developed in (III) will allow us to accurately determine the percentage that represents the insurer's retention. To further explanation, we review De Finetti's results:

$$\Theta i = \min \left[1, \max\ (0, 1 - \frac{\lambda\mu E(Xi)}{Var\ (Xi)})\right]$$

These results could be found based on results of nonlinear optimization that generalize the notion of Lagrange multipliers under inequality constraints.

[4] C.Hess "Actuarial methods of life insurance"; economica edition; France ; 2000, pp 16 .

In the general framework, the standard formulation for a non-linear problem is as follows:

In the general framework, the standard formulation for a non-linear problem is as follows:

$$\text{Min } f(x)$$

$$\text{S/C} \begin{cases} g_i(x) \leq c_j & j = 1, \ldots, m. \\ h_k(x) = d_k & k = 1, \ldots, p. \end{cases}$$

X ($x_1, \ldots, x_n$) is a vector of unknowns to be determined, the functions g_i and h_k are defined on R ^ n.

The introduction of Lagrangian allows to write:

$$L(x, \mu, \lambda) = f(x) + \sum_{j=1}^{m} \mu_j (g_i(x) - c_j) + \sum_{k=1}^{p} \lambda_k (h_k(x) - d_k)$$

Where μ_j (j = 1, ..., m) and λ_k (k = 1, ..., p) are Lagrange multipliers for the constraints $g_i \leq c_j$ and $h_k = d_k$ respectively. First-order conditions, known as Karush-Kuhn-Tucker conditions, are required for vector X optimization (for more demonstration in F. Glineur y and J.-F. Walhin work)[5].

V. DATA APPLICATION

For digital application, we used data from CAARAMA insurance campany in Algiers. Our work is presented as follow:

- The probability of ruin

Applying the probability of ruin to different reinsurance treaties would be a bit difficult, especially for non-proportional treaties that require more statistics. For our study, we will apply this theorem in a quota proportional treaty. This treaty is the most answered in Algeria given the lack of experience and control of risk in personal insurance.

To be able to calculate the probability of ruin for this portfolio, it must be verified that:

- ✓ Claims are independent and identically distributed
- ✓ The X_i which represents the disaster attributed to each police, follow the normal law.

It is already known that the claims are independent and identically distributed, remains to verify their normality :

TABLE 01: NORMALITY TEST RESULTS

```
data:  Charge de sinistre
    W = 0.64228, p-value < 0.22
```

Fig. 01 insurance campany CAARAMA data elaborated by R3.5

[5] F.Glineur y and J. F. Walhin; op.cit; pp 07.

If the p-value is below the fixed alpha level (often 0.05) then the null hypothesis is rejected and it is concluded that the disaster distribution is not normally distributed.

The p-value is equal to 0.22> 0.05, which implies the normality of the amounts of claims.

As a result, we continue:

$$E(X) = \frac{1}{n} \sum_i^n Xi = 3\,363\,088{,}08$$

$$Б(X) = \frac{1}{n} \left(\sum_i^n (Xi - E(X))\right) = 4\,178\,268{,}38$$

For a security load λ=15%

$$\Pi(X) = (1+\lambda)\, E(X) = 3\,867\,551{,}29$$

$$T = \frac{(10000000 + (0.15 * 3\,363\,088{,}08)}{4\,178\,268{,}38}$$

$$T = 2.51$$

We note that our safety factor is equal to 2.51 < 4 hence the need to resort to reinsurance. In this case, the retention interval is :

$$T(\theta) \geq 4 \Rightarrow \frac{R + \theta \lambda E(X)}{\theta \sigma(X)} \geq 4$$

$$\Theta < \frac{R}{(4\sigma(X) - \lambda E(X))}$$

$$\Theta < \frac{10000000}{((4 * 4\,178\,268{,}38) - (0.15 * 3\,363\,088{,}08)}$$

$$\Theta < 61.7\%$$

And so the probability of ruin is :

$$P(\text{ruined}) = 1 - F_0 \left(\frac{R + \theta \lambda E(X)}{\theta \sigma(X)}\right)$$

$$P(\text{ruined}) = 1 - F_0 \quad (4)$$

$$= 1 - 0.999968$$

P (ruined) = 0.0032%

We can confirm that reinsurance in fact reduce the probability of ruin of the insurance.

After having determined the interval of the retention which corresponds to the commitments of the insurer, one tries to find the percentage which makes it possible to optimize this retention thanks to the modeling of "De Finetti".

- Optimization by the De Finetti's modeling

De Finetti proposed to analyze proportional reinsurance structures by minimizing the variance of the insurer's gain under the constraint that the expected gain is fixed a priori.

To carry out this optimization process, we assume the following assumptions:

- The overall claim burden over a period of one year is: X= $\sum_{i=1}^{n} Xi$
- Each Xi represents the victim attributed to each police. For i = 1, ..., n, the Xi are independent and identically distributed
- The pure premium used to cover on average the loss ratio is worth P = E (X)
- The insurer applies an equal load to the identical pure premium for each policy.

This shipment is supposed to represent the profit of the insurer. We do not include in this percentage management and acquisition fees, as well as any taxes.

- The technical premium that can be used to pay the claims is therefore P=(1+ξ_i^{re})E(X)
- Optimal retention refers to θ.

We chose to solve this problem through the R software.

The latter allowed us to find an optimal value (or vector) (maximum or minimum) for a formula that is the objective to be achieved, under constraints or limits applied to the values or vectors sought, and this in order to solve linear or nonlinear problems, but it is limited to a reduced number of constraints. In the case of our problem, the size of our sample is very large, and as the number of constraints follows the number of variables, we face a large problem. But we can always find a workable solution for the reduced problem, while respecting the constraints.

The sub-portfolio we have selected includes the largest clients, representing 20% of the initial portfolio.

TABLE II : PREMIUM AND CLAIMS DATA sub-portfolio

insured	Pi	Xi
1	18 874 721,62	24 427 493,71
2	625 158,00	4 141 666,66
3	1 524 649,50	6 793 190,00
4	5 968 415,15	2 586 010,77
5	2 042 545,64	27 681 666,66
6	1 321 601,45	1 800 000,00
7	38 758 259,58	941 666,66
8	54 145 614,51	2 278 125,00
9	28 293 239,63	941 666,66
10	87 852,76	600 000,00
11	124 794,00	100 000,00
12	9 426 211,76	907 257,23
13	686 164,67	981 531,44
14	691 900,00	516 000,00
15	344 824,64	2 051 648,75
∑	162 5 952,92	76747 923,54

Fig. 02 insurance campany CAARAMA data elaborated by R3.5

It should be noted that this sub-portfolio is rather homogeneous, insofar as premiums collected by the company sometimes cover the amounts of claims and make a profit, other times they take losses. But the turnover is still positive. However, the loss ratio is still important, and therefore will affect, possibly, the retention of the insurer.

To determine the average profit, we used the result found by the ruin probability method, that is, calculate E (Z (61.7%)).

E(Z(61.7%))= 162 915 952.92-(0.15*0.617*10 603 839.13)-10 603 839.13

E(Z(61.7%))= 151 702 923,2 DA

We can conclude that this amount represents the maximum expected profit that an insurer could expect. Given its reserves which have limited the level of retention to 61.7%, the insurer is not aiming for more than its expected profit, although it can increase its real profit if it keeps more than 61.7%, but it does not have the necessary reserves which would allow it, in the event of the occurrence of disasters, to assume its retention.

We can also calculate the minimum expected profit to make it vary between the two terminals:

E(Z(0%))= 152 368 128,8 DA

On R 3.5 we have varied the amount of profit expected to observe, for each amount, the level of retention that minimizes the volatility of claims. And this, proceeding as follows:

```
lp (direction = "min", objective.in,
const.mat, const.dir, const.rhs,
        transpose.constraints = TRUE,
int.vec, presolve=0, compute.sens=0,
        binary.vec, all.int=FALSE,
all.bin=FALSE, scale = 196, dense.const,
        num.bin.solns=1, use.rw=FALSE)
```

Our results are as follows:

TABLE.III: RENTENTION RESULTS

The expected gain	The reinsurer loading	The retention	Min σ(X)
153 217 000,00	15%	61,70%	2 643 006,26
153 150 000,00	15%	56.83%	2 242 253,10
153 050 000.00	15%	49.56%	1 705 372,60
152 950 000,00	15%	42.29%	1 241 847,81
152 890 000,00	15%	37.93%	998 944,21
152 800 000,00	15%	31.39%	648 105,44

Fig. 03 insurance campany CAARAMA data elaborated by R3.5

For each expected return value, R application has found a level of retention considered optimal because it minimizes the loss volatility.

The more the expected profit increases, the more the retention of the insurance will be important, but this is not reassuring because the volatility of the accident increases with the expected return value. Indeed, a company wishing to improve the profitability of its portfolio must accept to bear more risks.

The volatility of this portfolio is not really significant compared to the expected profit. Here it is a portfolio that contains good risks. However, it is important to consider the level of retention that minimizes the variance in earnings.

The volatility of the claim provides a good indication of the risk of loss. Indeed, when a company decides to increase its level of retention to aim for a more attractive profit, it must expect a greater loss volatility that could hit its portfolio and fluctuate its results.

The company CAARAMA insurance has opted for a retention of 50% applied to this portfolio. Given the lack of data, we can assume two cases:

- ✓ The expected profit was DA 153,050,000;
- ✓ Or the negotiations with its reinsurers ended up agreeing on 50% as retention.

This level of retention can be considered optimal because it reduces the volatility of the claim, but only if the expected profit is equal to 153 050 000 DA.

Determining the optimal level of retention for this portfolio depends on the policy of the insurer and its attitude to risk (which is measured by the volatility of the claim), the insurance company must target a goal rather realistic that corresponds to its target of the year and optimize it.

The level of optimal retention will not necessarily be put into practice in the portfolio concerned, it will also depend on negotiations with the reinsurer, which in turn optimizes its own retention.

VI. CONCLUSION

Nonlinear optimization provides a key decision-making tool in quantitative finance, and our work is a critical illustration of the importance of convex function studies in decision-making.

As for the insurer, the principle of risk sharing between him and reinsurer is based on actuarial methods that determine the commitment that corresponds to each party. For the retention to be optimal, it must correspond to the strategies predefined by the insurer according to the expected return torque, loss volatility.

The optimization process can be achieved by combining two methods that aim respectively at determining the retention interval and optimizing it for a fixed expected benefit.

The theory of ruin, helped us to:

- ✓ Find the interval of our retention, which should be less than 61.7%;
- ✓ To note the positive effect of reinsurance on the insurer's security by minimizing its probability of ruin;
- ✓ To calculate the maximum expected profit of the insurer, which was subsequently used in the "De Finetti" method.

Lastly, the application of the "De Finetti" method allowed us to determine the exact value of the retention, which contributes to the profit desired by the insurer, and for which the volatility of the claims is minimal. But this value is not necessarily put into practice, everything depends on the negotiations with the reinsurer, and the attitude of the insurer against the risk.

We can propose as extension for our study a realistic modeling of the expected profit value.

REFERENCES

[1] C.Hess "Actuarial methods of life insurance"; economica edition; France ; 2000, pp 16-22.

[2] F. Glineur y and J. F. Walhin; "Finetti's Retention Problem for Proportional Reinsurance Revisited", Secura, Belgium, 2006., pp.02–08.

[3] J. Blondeau C.Partrat.; "Reinsurance: a technical approach"; economica edition; France, 2003, pp 14-83.

[4] P. Constantin Niculescu, L.E Persson ; "Convex function and their application, a contemporary approach", Springer edition, New York, 2004, pp145-157.

[5] P. Azcue and N. Muler; "Optimal reinsurance and dividend distribution policies in the Cramér–Lundberg model", Math. Finance 15, 261–308.

.

A vehicular network architecture based on Internet of Vehicles for improving the urban traffic management

Somia BOUBEDRA
LRS Laboratory, Computer Science Department
Badji Mokhtar University
Annaba, Algeria
as_boubedra@esi.dz

Cherif TOLBA
LRS Laboratory, Computer Science Department
Badji Mokhtar University
Annaba, Algeria
ctolba@yahoo.fr

***Abstract*—we propose a novel architecture based on Internet of Things for managing and monitoring urban traffic system, our proposition is based on several technologies, selected carefully, such as wireless sensor network, RFID radio identification technology, Fog/Edge computing, and Cloud Computing. We have used and combined all of these technologies to lead to the big problems of the road traffic such as congestions, accidents, and pollutions.**

Keywords— Internet Of Things (IoT); Internet Of Vehicles (IoV), urban traffic system; Fog/Edge servers, Sensors, RFID, Master Nodes.

I. INTRODUCTION

In our daily life, the technology becomes an important aspect, as it plays a major role in all domains, and offers great benefits to individuals and societies. This involves the exponential growing of the concept of Internet of Things, which is the interconnection of billions of different types of devices, and sensors, called "smart objects", so, they cooperate to meet our needs, with restricted capacities in terms of energy, memory, and processing powers [1].

Moreover, Transportation systems have a strong impact on the development of our society. Effective movement of goods and people contributes to economic growth and changes our territories through a good accessibility. That is why the development in transportation is one of important factors to indicate the well-being of a country [7]. In addition, the use of New Information Technologies and Communications to improve the transportation systems become a central solution in the field. The increase in computing power and the great development of the embedded systems, as well as the quality of sophisticated sensors, have made it possible to propose more effective control mechanisms; and better consideration of operators or users, the result is so-called Intelligent Transportation Systems (ITS).

However, the European report [2] on the evaluation of research programs in transport, Intelligent Transport Systems are considered vital for designing sustainable transport systems. According to this report, through the integration of information, communication and control technologies, ITS enable authorities, operators and individuals to make better decisions. ITS concern all systems that improve the use of means of transport using a set of technologies to meet the objectives of the domain.

Whatever the functionality associated with the ITS, it is built from data captured on the network, which is received and processed by software. As a result, all the advances in communications, sensors and computing are potentially benefiting the transportation systems. For example, the development of connected or autonomous vehicles is only possible through the implementation of communications between vehicles and with a suitable infrastructure, the deployment of high-performance sensors, and significant computing capabilities.

In addition, researchers of urban traffic systems have oriented their researches to the use of the Internet of Things' technologies, which led to the apparition of new concept: The Internet of Vehicles. IoV is based on the Internet, wireless sensor networks and sensing technologies to perform both intelligent recognition of road users (who are considered as objects), monitoring, and finally the management and the real-time treatment of road traffic.

To discuss the details of this topic, we have organized the rest of our paper as following: in the next part, we explain the problems and motivations, which led us to work on this area, then, we present general notions about the topic, by defining the Internet of Things, the internet of vehicles, and the urban traffic system. Section IV is devoted to present the IoT architectures' background. Next, we pass to illustrate our proposed architecture to overcome the problems of the urban road traffic mentioned in section II. Finally, we present the conclusion, and the perspectives of this research work.

II. PROBLEM STATEMENT

In recent years, the large number of vehicles has led to a considerable increase in urban traffic. As a result, road traffic has become one of the major problems in most major cities. Road traffic problems are congestions and accidents resulting huge loss of time, damage to property and environmental pollution. In addition, according to [12], a report from "Automotive News", states that the number of cars connected to the Internet worldwide will increase to 152 million in 2020.

These issues explain why many research programs around the world aim to improve our transportation systems; this is indeed a difficult task because the distributed, open, dynamic and partially controllable nature of transport networks makes it a complex area.

This research work deals with the problem of managing and monitoring an intelligent transportation system, especially the urban traffic system, the aim of our contribution is limiting the nuisance caused by the increase in the use of transport. Thus, better mobility means limiting the environmental impact of the pollution generated, and improving safety and conditions of people's life.

III. General Notions

A. Internet of things IoT

In [13], IoT was defined as a "dynamic global network infrastructure with self-configuring capabilities based on standards and interoperable communication protocols; physical and virtual 'things' in an IoT have identities and attributes and are capable of using intelligent interfaces and being integrated as an information network".

From the viewpoint of network, the IoT is a very complicated heterogeneous network, which includes the connection between various types of networks through various communication technologies [8].

In addition, the Oxford Dictionaries offers a concise definition of the IoT: Internet of things (noun): The interconnection via the Internet of computing devices embedded in everyday objects, enabling them to send and receive data [9]

Furthermore, the capabilities offered by the IoT can save people and organizations time and money as well as help improve decision-making and outcomes in a wide range of application areas.[10]

As well, IoT plays an important role in transportation field, such, vehicles have increasingly powerful sensing, networking, and data processing capabilities For instance, IoT technologies make it possible to track each vehicle' existing location, monitor its movement, and predict its future location. [8]

B. Internet Of vehicles IoV

It is a dynamic network, which consists of IoT enabled cars by using modern embedded and electronic devices like sensors and GPS, and integration of the information and communication systems to improve traffic flow, and to offer more effective road management and accident avoidance.

The urban traffic system has benefited from a lot of IoT applications like 'Internet of Vehicle' concept, Vehicle-to-Vehicle (V2V), and Vehicle to Infrastructure (V2I) communications, and have been transformed to a new level of interoperability, stability and efficiency, because, If vehicles communicate with each other, risks for accidents and mishaps would be very low. In addition, by using IoT technologies in the road traffic, we can monitor urban transportation systems, determine the state of traffic and pedestrian densities, identify damages and accidents, avoid collisions as needed, and optimize travel route [6].

IV. IoT Architectures' Background

In [4], authors surveyed existing IoT architectures, which are three-layer architecture, Middleware-based architecture, Service Oriented Architecture (SOA), and Five-layer architecture. Furthermore, they marked that the five-layer architecture is the most appropriate model for IoT applications, due to its simplicity, by the way, this later consists of five layers : 1. Objects layer or perception layer, which contains physical components like sensors, actuators, 2. Objects Abstraction layer, by using this layer we transfer data generated by Objects layer over WiFi, GSM... 3. Service management layer, which processes data, makes decisions, and delivers services over network protocols. 4. Application layer, that provides high quality smart services to meet customer's needs, and 5. The Business layer that supports decision-making based on big-data analysis.

Authors of [5] presented two types of IoT architectures; the basic *Three-layer architecture*, it consists of perception or sensor layer, Network layer and application layer, and the four-layer SoA-based IoT architecture, which is composed of Perception layer, Network layer, Service layer, and Application layer, service layer is made of service discovery, service composition, service management, and interfaces. According to the authors, the service-oriented architecture is more flexible and generic, because a service layer is developed between network layer and application layer to provide the data services in IoT architectures like data aggregation and processing in network layer, and data mining, data analytics in application layer. After that, they introduced the relevant enabling technologies and challenges of each layer, and they token the four-layer SoA-based IoT architecture as an example.

In [3], authors proposed a four-layer architecture for future heterogeneous IoT, which contains Sensing layer, Networking layer, Cloud computing, and Application layer, we explain each layer with more details in section V.

V. METHODOLOGY

In [3], authors propose a four-layer architecture for future Internet of Things; we combine this architecture by the concept of fog/edge computing, we add a novel layer to this architecture, which is the edge servers' layer. Then we adapt this architecture to be destined to the road traffic systems; in this section, we explain our architecture with more details.

A. Layers of the proposed architecture

First, we present the layers of our architecture illustrated in figure 1:

- **Sensing layer:**

This layer represents the physical sensors, actuators and RFID tags that aim to capture, collect, and transmit information [4].

A large number of sensors are deployed in the monitoring area [3], which is in our case the urban road; we use sensors to collect data about the state of the road (if it is congested, or there is an accident or a fire in the road). From vehicles that are equipped by RFID tags, and pedestrians who have all smartphones in their possession, or swatches connected to the internet.

Those sensors send the captured data to the sink node, which we call the master node; we will explain its role in the Fog/Edge computing layer.

- **Network layer:**

In this layer, we implement network protocols, and the corresponding topologies like star topology, tree topology, mesh topology, or hybrid topology, in order to forward data packets from source node to destination node [3];

However, we consider self-organizing network protocols, because we need more robustness and efficiency in construction of network topology, like the IPv6 routing protocol "RPL", which is a distance vector routing protocol designed by the Internet Engineering Task Force IETF, for Low Power and Lossy Networks.

- **Cloud layer:**

This layer is very important to handle the tremendous amount of data collected, and transmitted by other layers to cloud servers and big data centers, to be processed, stored, and to make decisions based on data analysis [3][5], thanks to the powerful analytical computing capacities that have cloud servers.

Cloud computing is now a mature technology used to create, store, and use data over the Internet. Although, when a massive amount of data need to be stored, processed, and analysed efficiently in data centers and cloud servers , a new technology appear to fulfil the gap, which is the Fog/Edge computing, to extend cloud computing to be closer to the network of the things [5].

- **Satellite Sub-layer:**

To transmit data between Edge Servers and Cloud data centers and servers, we use Satellites, to gain time, throughput, and energy.

- **Fog/Edge Computing layer:**

In this layer, we have two types of devices, the Master Nodes, and the Edge servers.

We can use Edge servers for insuring processing, storage and make decisions near to the network, instead of doing all the computations in the cloud servers, hence, Edge computing have faster response and greater quality than cloud computing [5], especially, when we are faced to a real time application like road traffic.

We update data centers of the cloud once a day, in the night; on the other hand, we transmit data from master nodes to Edge Servers several times and periodically in the journey, because we can place some types of data for further computations and analysis, however, the high priority data, we address it immediately to the closest Edge server, to insure the real time property of the road traffic system.

The master node is an access point with a good processing, energy, and transmission capacities, if we compare it with the road sensors, its role is to 1- receive the data collected by all the sensors near of it, i.e. In the same area. 2- After that, it makes some calculations and data aggregations to reduce the big amount of the collected data, because and without a doubt, we will find lot of redundancy, because, the sensors are in the same region and they will capture sometimes the same information.

- **Application layer:**

The application layer responds to users' needs, by providing them the corresponding services [4], for instance, a car driver needs to know if this road is congested or not, he uses our application to get the best response. Our application here is the urban road traffic management, which includes vehicles, pedestrians with their smartphones, or smart watches, road sensors, and other smart devices are connected as objects in the network, delivered data is used to insure the real time management of the urban road traffic. Edge and cloud servers can manage and monitor remotely the objects based on data analytics and visualisation [3].

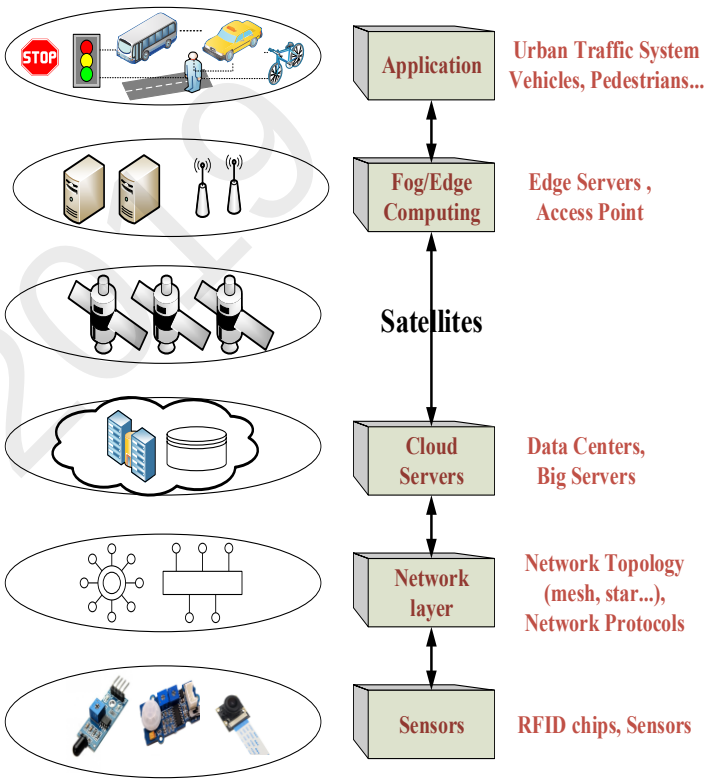


Fig 1. The proposed architecture's layers

B. The proposed architecture

Our Architecture is ***hybrid***, in terms of connecting objects in the IoT network, it means that objects ***cooperate*** with each other and exchange information of traffic; and ***hierarchical***, because objects connect with the master node to transfer the data captured by sensors to the Edge Servers, in addition, data processed in edge servers will be send for processing and make decisions to the cloud centres, as it's illustrated in figure 2.

In each vehicle, we find a *GPS* (Global Positioning System); this later is responsible for receiving important data like location, time and weather condition from satellites [7]. In addition, we have *RFID* chips; its role is to exchange information with other vehicles and pedestrians and with road sensors using *Zigbee IEEE 802.15.4*.

Road sensors are responsible for capturing road traffic data, from vehicles and pedestrians, these data tell us if there are congestions, accidents, flames…, after that, they transmit the collecting data to the master node via *WiFi IEEE 802.11.*

The master node consists of a communication and data treatment modules; the communication part is a wireless antenna, which is responsible for receiving and decoding the transmitted data packets from the road sensors or the edge servers. Furthermore, the data treatment module is used to do some data aggregation on the data received from the road sensors, because there will be certainly redundancies, in addition, mechanisms of data aggregation aim to reduce the amount of transmission data and the energy consumption [7].

The aggregated data are forwarded to the nearest edge server via GPRS (General Packet Radio Service), which is a cellular communication protocol, named as 2.5 Generation (2.5 G), it means that is between the second generation and the third generation of GSM (Global System for Mobile communication).

Each Edge Server make processing and calculations on data transmitted from Master nodes, make decisions, and preventions, to raise alarm to drivers or pedestrians to avert them if there is congestions, accidents, flames… to avoid more damages in the road; this process is repeating during all day long.

Edge servers have a big power of storage and processing to make better decisions to ameliorate the quality of transportation in urban areas, they are an intermediary between Cloud servers and data centres, and sensor networks in the road.

Processed data, decision make, and preventions will be sent to Cloud servers through satellites, we use 4G to transfer data.

Why using these existing protocols? We use any available network within the range, for insuring communication between components in an IoT system, seems to be a better solution [7]. Like here in our case, we use WiFi, and cellular networks like GPRS and 4G LTE, which are a pre-existing network architectures, in order to avoid implementing new infrastructures.

In Cloud centres, we make global and heavy operations, due to the big capacity of processing and storage, we use virtualisation and data analytics to make better decisions and preventions and store data to use it for improving the urban road traffic system.

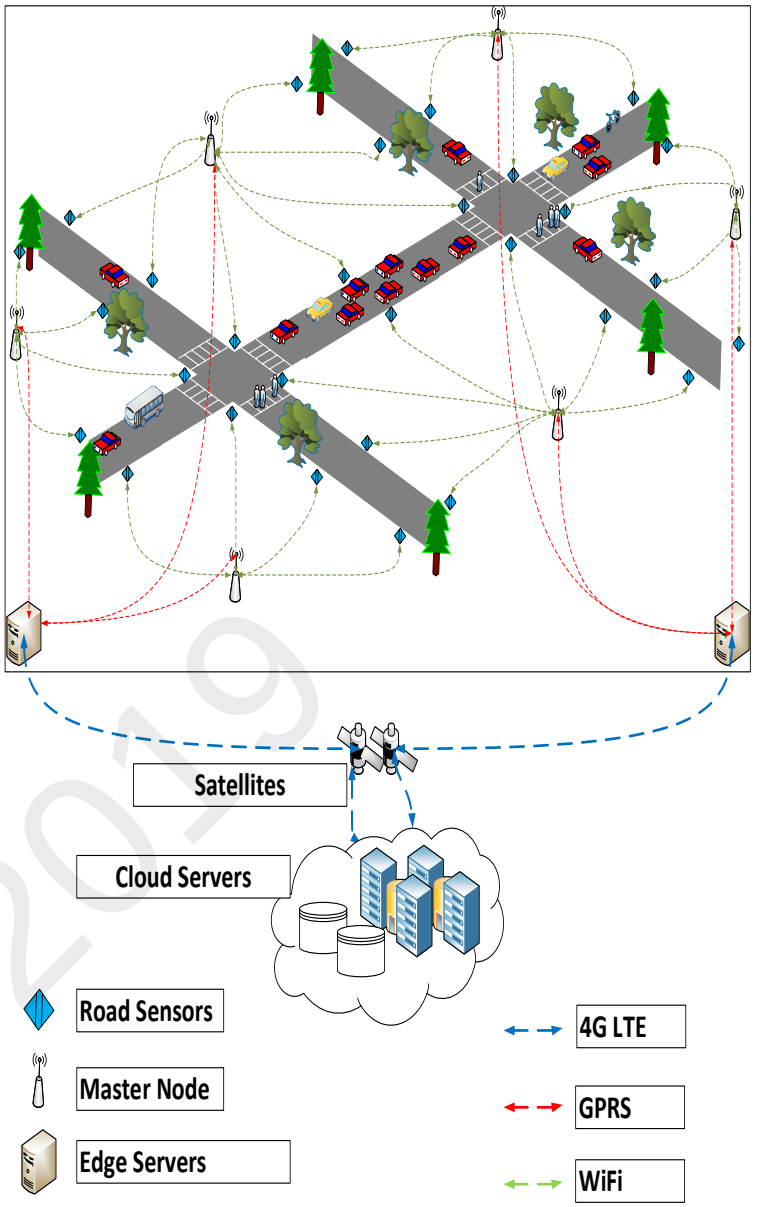


Fig 2. The proposed architecture of the urban traffic system based IoV

CONCLUSION & PERSPECTIVES

In the present article, we present a solution for the management of the urban road traffic system; we propose a Five-layer architecture, based on the scientific work of [3]. Therefore, the proposed architecture is generic and flexible for all the urban traffic systems, and can be applicable in the real word, because we bring together current and existing IoV and IoT technologies.

The proposed architecture is global; we work to detail it more and more, using IoT and IoV technologies, and to implement its layers in the near future; once successfully implemented, the reduction of damages, collisions, congestions, and pollutions in the urban road traffic will certainly benefit the quality of people's life in urban areas.

REFERENCES

[1] J. Subramaniam, L. H. Yean, and S. Manickam, "Intelligent IPv6 based IoT network monitoring and alerting system on Cooja framework," in Journal of Fundamental and Applied Sciences, vol. 9, pp. 661-670, 2017.

[2] "Innovate for a competitive and economical resources of transportation systems", Technical Report - Europeen Union, 2012.

[3] T. Qiu, N. Chen, K. Li, M. Atiquzzaman, W. Zhao, "How Can Heterogeneous Internet of Things Build our Future: A Survey", in IEEE Communications Surveys and Tutorials, vol. 20, no. 3, pp. 2011-2027, 2018.

[4] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications", in IEEE Communication Surveys & Tutorials, vol. 17, no. 4, pp. 2347-2376, 2015.

[5] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications", in IEEE Internet of Things Journal, vol. 4, no. 5, pp. 1125-1142, 2017.

[6] P. Suresh, V. Daniel, V. Parthasarathy, R.H. Aswathy, "A state of the art review on the Internet of Things (IoT), History, Technology and fields of deployment", in International Conference on Science, Engineering and Management Research (ICSEMR 2014), Chennai, India, 2014.

[7] S. Vongsingthong, and S. Smanchat, "Internet of things- a review of applications & technologies", in Suranaree Journal of Science and Technology, vol. 21, no. 4, pp. 359-374, 2014.

[8] L. D. Xu, W. He, and S. Li, "Internet of Things in Industries: A Survey", in IEEE Transactions on Industrial Informatics, vol. 10, no. 4, pp. 2233-2243, 2014.

[9] K. Rose, S. Eldridge, L. Chapin, "The Internet of Things: An Overview - Understanding the Issues and Challenges of a More Connected World", Internet Society, 2015.

[10] A. Whitmore, A. Agarwal, and L. D. Xu, "The Internet of Things - A survey of topics and trends", in Information Systems Frontiers, vol. 17, no. 2, pp. 261-274, 2015.

[11] P. P. Ray, "A survey on Internet of Things architectures", in Journal of King Saud University- Computer and Information Sciences, vol. 30, no. 3, pp. 291-319, 2018.

[12] E. Ahmed, I. Yaqoob, A. Gani, M. Imran, and M. Guizani, "Internet-of-Things Based Smart Environments: State of the Art, Taxonomy, and Open Research Challenges", in IEEE Wireless Communications, vol. 23, no. 5, pp. 10-16, 2016.

[13] S. Li, L. D. Xu, and S. Zhao, "The internet of things: a survey", in Information Systems Frontiers, vol. 17, no. 2, pp. 243–259, 2015.

A cooperative-based approach towards fully autonomous driving with consideration of control uncertainty

Oussama MESSAOUDI
LaSTIC Laboratory, Department of Computer Science
University of Batna 2
Batna, Algeria
oussama.messaoudi@univ-batna2.dz

Ammar LAHLOUHI
LaSTIC Laboratory, Department of Computer Science
University of Batna 2
Batna, Algeria
ammar.lahlouhi@gmail.com

***Abstract*—The automation of the driving task in a stochastic and partially observable environment is currently incomplete due to control and sensing uncertainty. Many approaches were proposed in an attempt to reduce sensing uncertainty, whereas control uncertainty was completely ignored. In this paper, we propose a cooperative-based approach in order to handle only control uncertainty in a stochastic and fully observable environment. Also in this paper, we focus on the automation of only a part of the driving task, in particular the automation of the longitudinal control task.**

Keywords— agent-based modeling and simulation; autonomous driving; artificial intelligence; longitudinal control; microscopic simulation; uncertainty;

I. INTRODUCTION

During the last two decades, driving assistance systems have significantly reinforced road safety and helped to provide a comfortable and enjoyable driving experience. However, with the continuous increase in the number of road users nowadays, the need for a safer and more efficient driving becomes more and more important. The current technological advancement can provide one viable solution to answer such needs, the implementation of autonomous driving systems. However, the development and deployment of these systems still face many challenges, one major challenge is how to resolve control and sensing uncertainty.

An environment is said to be uncertain if it is stochastic and/or partially observable [1]. Also, uncertainty originates from the inaccuracy of control actions, non-modeled external influences and the use of partial or noisy information about the state of the environment [2].

In the case of the autonomous driving system, sensing uncertainty originates from a partial, incomplete and/or noisy perception of the surrounding environment. In the last decade, many autonomous driving systems addressed sensing uncertainty to provide an accurate estimation of the autonomous vehicle's surroundings [3-5]. This should allow the autonomous vehicle to produce a reliable and trusted autonomous driving in a partially observable environment.

On the other hand, control uncertainty originates from inaccuracy of the vehicle's control actions, which is caused by tire slip, wheel spin and/or by brake lock-up. As a result, the environment in this case, more pacifically the real world, is considered to be stochastic [6, 7].

In the past, many systems ignored control uncertainty to autonomously drive the vehicle in an environment considered as deterministic and uncertainty-free (e.g. see [3, 8, 9]). Such systems could cause collisions if deployed in a stochastic and uncertain environment due to ignoring the inaccuracy of the vehicle's control actions.

To this end, it is crucial to handle control uncertainty in order to produce and maintain completely autonomous, reliable and collision-free driving.

The human driver has an intelligent behavior allowing him to drive in various challenging, difficult and continuously changing environmental conditions. In fact, his decision process maps inputs (estimated percepts) to outputs (actions) using an unknown reward function and an unknown transition model. This allows the human to estimate the inaccuracy of the vehicle's actions as well as the state of the surrounding traffic, and as a result, handle both control and sensing uncertainty.

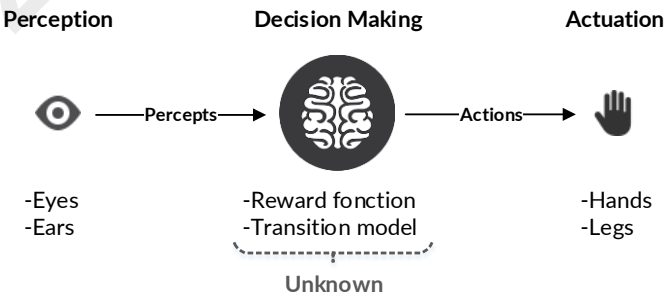


Fig. 1. The human driver's decision process.

Therefore, instead of completely eliminating the human driver from the control loop, why not exploit and benefit from his resources and skills to handle uncertainty. One possible approach to exploit the human driver's resources and skills is through the cooperation with the autonomous driving system.

In this paper, we propose a cooperative-based approach allowing an autonomous driver agent to cooperate with and learn from the human driver before taking over the driving task. Also, we focus in this paper on handling control uncertainty and also on automating the longitudinal (velocity) control task.

II. SYSTEM OVERVIEW

A. System architecture

The main contribution of the work presented in this paper is to integrate the human driver in the control loop in order to cooperate with the autonomous longitudinal control system. This should allow our system to exploit the human driver's resources and skills, to take the inaccuracy of the vehicle's actions into consideration and to eventually handle control uncertainty.

In MASA-Method, a multi-agent system (MAS) modeling methodology proposed in [10], an agent H consists of the human operator and an agent F equipped with an interface. The F agent envelops and represents the human operator in the multi-agent system as follows:

- In response to other agents' requests (software agents', robots' of other humans'), the F agent asks the human operator to execute certain procedures in order to accomplish a given task,
- The F agent passes the human driver's requests to other agents of the MAS and asks them to perform certain actions destined to them,
- Also, through the F agent's interface, an agent of the MAS can pass any information to the human operator (e.g. system state).

As a result, this approach allows different agents of the multi-agent system to interact, communicate and cooperate with the human operator. Furthermore, different agents of the MAS can benefit from the human operator's resources and skills in order to accomplish their tasks.

Therefore, in order to exploit the human driver's resources and skills and handle control uncertainty to eventually automate the longitudinal control task, we propose to use MASA-Method as follows:

- We integrate the human driver in the control loop to make his own decisions and manually drive the vehicle while taking control uncertainty into consideration,
- In order to control the velocity of the vehicle, the human driver acts on a human-machine interface, more specifically the interface of the F agent,
- The F agent perceives the human driver's actions and passes them to an autonomous agent (the $DAgent$) for execution,
- In addition to executing the human driver's actions, the $DAgent$ communicates the vehicle's state (e.g. velocity, leading vehicle's velocity, gap to the leader, etc.) with human driver through the interface of the F agent.

To this point, the proposed autonomous longitudinal control system consists of:

- ***The H agent:*** this agents represents the human driver in the autonomous longitudinal control system. More specifically, this agent consists of the human driver and the F agent.
- ***The $DAgent$*:** this agent is responsible of exploiting the resources shared by the H agent, more specifically those of the human driver's.

More specifically, the role of the $DAgent$ is to learn from the human driver's behavior and eventually provide autonomous and collision-free longitudinal control in a stochastic and uncertain environment.

B. Learning from the human driver's behavior

Before we tackle the challenge of learning from the human driver's behavior, we first introduce two definitions important to the understanding of the problem at hand.

First, an environment's transition model describes all the possible outcomes of performing any given action in any given state. Second, the reward function gives an agent a reward for achieving a specific state s'_t at time step t. This reward can also depend on the initial state $s_{t-\Delta t}$ and the action $a_{t-\Delta t}$ performed at $t-\Delta t$. Furthermore, the reward values are used for decision making under control uncertainty. More specifically, they are used to calculate the expected utility value of performing any given action a in a certain system state s, and then to identify an optimal action that yields the highest expected utility [1].

The human driver has a very intelligent and complex behavior that allows him to drive his vehicle in various challenging, difficult and continuously changing environmental conditions. In fact, the human driver accomplishes his task based on:

- ***An unknown transition model:*** the human driver is capable of establishing an accurate estimation of the inaccuracy of the vehicle's actions under any given condition (e.g. on a dry, wet, snow or icy road).
- ***An unknown reward function:*** based on his estimation of the inaccuracy of the vehicle's actions, the human driver always chooses an action that yields the highest expected utility. More specifically, such action should maintain collision-free driving with consideration of control uncertainty.

To this end, we can say that the human driver uses an implicit process to measure the rewards of its actions, and uses also a hidden transition model to estimate the inaccuracy of his actions and handle control uncertainty.

The major challenge of automating any given task under control uncertainty, including the driving task, is that both the environment's transition model and the human operator's transition model are unknown.

Therefore, by observing the human driver's behavior and exploiting his resources and skills, the autonomous agent can learn to handle control uncertainty and to drive the same way the human does. As a result, the $DAgent$ must:

- Learn the environment's transition model which allows our system to (i) achieve an accurate estimation of the inaccuracy of its accelerations and, therefore, (ii) take control uncertainty into consideration.
- Learn the human driver's reward function in order to (i) learn how to measure the expected utility values of its actions and (ii) learn how to choose each step an action that yields the highest expected utility.

In this paper, we focus only on learning the environment's transition model in order to handle control uncertainty and provide autonomous and collision-free longitudinal control.

Also in this paper, (i) we highlight the importance of building an accurate estimation of the environment's state, (ii) we identify the challenges of building an accurate transition model and (iii) we examine the effects of using an inaccurate transition model.

C. Simulation of the human driver behavior

Before we tackle the approach proposed to handle control uncertainty, it should be noted that the behavior of the human driver, more specifically his decision making process, is simulated using the optimal velocity robust car-following model ($OV-RCFM$) proposed in [11] (see (1)).

$$\pi^*(s) = \arg\max_{a\in A} \min_{s'\in T_s^a} R(a,s,s') \quad (1)$$

The $OV-RCFM$ model is an extended version of the original $RCFM$ model proposed in [12]. Furthermore, the $OV-RCFM$ is a car following model capable of providing collision-free longitudinal control in various configurations of an uncertain simulation environment. To accomplish its task, this model uses a reward function R (see Algorithm 1) in addition to a pre-built and accurate transition model T.

ALGORITHM 1 THE REWARD FUNCTION USED IN THE $OV-RCFM$ MODEL (see table I for legend description).

Input: $a, s, s', v_l, g, t_r, T_s^{MaxDecel}, v_{max}, g_{min}, MaxAccel, MaxDecel$
Output: u: real

Initialization
$v_{free} = s + MaxAccel \times \Delta t$
$g_e = \max\{0, g - g_{min}\}$
$b_e = \min_{s'\in T_s^{MaxDecel}} \frac{|s'-s|}{\Delta t}$
$vSafe = v_l + \frac{g_e - v_l t_r}{\frac{v_l+s}{2b_e} + t_r}$
$v_{des} = \min\{v_{max}, v_{free}, vSafe\}$

Compute the reward
if $s' \le v_{des}$ **then**
  if $s' = v_{des}$ **then**
    return 1
  else
    return $\frac{s'}{v_{des}}$
  end if
else
  return $\frac{v_{des}}{s'}$
end if

TABLE I. DESCRIPTION OF EACH LEGEND USED IN ALGORITHM 1.

Legend	Description
s	The vehicle' state at time step t. More specifically, s represents the vehicle's velocity
a	An action, more specifically the acceleration ($a \in [-Maxdecel, MaxAccel]$) under evaluation
s'	The vehicle's possible velocity at the next time step $t+\Delta t$
$T_s^{MaxDecel}$	All possible outcomes of applying $MaxDecel$ on s at t
g	The current gap to the preceding vehicle
g_{min}	The minimum gap the vehicle must respect
g_e	The effective gap to the preceding vehicle
$MaxAccel$	The vehicle's maximum acceleration
$MaxDecel$	The vehicle's maximum deceleration
b_e	The vehicle's maximum effective deceleration
t_r	The reaction time (by default, $t_r = 1\ sec$)
v_l	The leading vehicle's velocity
v_{max}	The following vehicle's maximum velocity
v_{free}	The free-flow velocity
$vSafe$	The safe velocity
$vDes$	The desired velocity

Therefore, in the system proposed in this paper, a software agent ($HAgent$) implements the $OV-RCFM$ model to take over the role of the human driver and to simulate his optimal longitudinal control behavior in a stochastic and uncertain environment.

D. Structure of the decision process

In this paper, the $HAgent$'s optimal longitudinal control behavior is based on the $OV-RCFM$ model and the reward function R. Also, in addition to learning the environment's transition model by observing the $HAgent$'s behavior, the $DAgent$ implements the same $OV-RCFM$ model and the same reward function R.

This allows us to focus only on learning the environment's transition model and examining the effects of its inaccuracy on handling control uncertainty during the autonomous longitudinal control.

To this end, our proposed cooperative longitudinal control system consists of two software agents:

- ***HAgent*:** is a software agent responsible for simulating the manual driving behavior and providing collision-free longitudinal control,
- ***DAgent*:** is an autonomous software agent responsible for learning the environment's transition model to eventually handle control uncertainty and provide autonomous longitudinal control,

In order to exploit the human driver's resources and skills, and to eventually achieve autonomous and collision-free longitudinal control in a stochastic and uncertain environment, we propose the following decision process:

- ***Stage 1: Exploration*** - building a transition model to handle control uncertainty,
- ***Stage 2: Exploitation*** – producing an autonomous longitudinal control with consideration of control uncertainty,

During the first stage, the autonomous agent ($DAgent$) observes the actions performed by the human agent ($HAgent$) and their outcomes to build the environment's transition model. To accomplish this task, more specifically after executing any action, the autonomous agent includes in the new transition model:

- The vehicle's velocity $s_{t-\Delta t} \in S$ at time step $t-\Delta t$,
- The action $a \in A$ applied at time step $t-\Delta t$,
- The outcome, more specifically the velocity $s_t \in S$ achieved at time step t after applying and acceleration a on a velocity $s_{t-\Delta t} \in S$ at $t-\Delta t$.

The new transition model should help the autonomous agent to later take into consideration the inaccuracy of the vehicle's accelerations during autonomous decision making.

Also in the first stage, while the $HAgent$ simulates manual driving, and while using the transition model that it is currently building the $DAgent$ makes decisions without executing them. In addition, the $DAgent$ continuously compares its decisions to the $HAgent$'s actions to calculate a ratio reflecting the similarity of both agents' decisions and actions.

The ratio value gives the human operator a feedback of the $DAgent$'s ability to take over and allows him to decide to or not to engage the autonomous longitudinal control. We present below four different formulas to calculate the ratio values; each formula allows for a specific margin of error as follows:

- $r_0 = \frac{n_{\Delta a=0}}{nT}$ is the number of actions $n_{\Delta a=0}$ that meet the condition $|\Delta a| = |a_{DAgent} - a_{HAgent}| = 0\ m/s^2$ divided by nT the total number of exercised actions. Here, the ratio value r_0 increases each time the decision made by the $DAgent$ matches the action exercised by the $HAgent$ ($\Delta a = 0\ m/s^2$).
- $r_{0.25} = \frac{n_{\Delta a=0.25}}{nT}$ is the number of actions $n_{\Delta a=0.25}$ that meet the condition $|\Delta a| \leq 0.25\ m/s^2$ divided by nT the total number of exercised actions.
- $r_{0.50} = \frac{n_{\Delta a=0.50}}{nT}$ is the number of actions $n_{\Delta a=0.50}$ that meet the condition $|\Delta a| \leq 0.50\ m/s^2$ divided by nT the total number of exercised actions.
- $r_{0.75} = \frac{n_{\Delta a=0.75}}{nT}$ is the number of actions $n_{\Delta a=0.75}$ that meet the condition $|\Delta a| \leq 0.75\ m/s^2$ divided by nT the total number of exercised actions.

Then in the second stage, and after the human operator decides to engage autonomous longitudinal control, the $DAgent$ implements the $OV - RCFM$ model, uses the reward function presented in Algorithm 1 and, most importantly, uses the learned transition model to handle control uncertainty and control the velocity of the vehicle.

To this end, in order to handle control uncertainty and only automate the longitudinal control task, we proposed to start by learning the environment's transition model based on observing the $HAgent$'s behavior who simulates the human driver's decision making.

III. NUMERICAL RESULTS

A. Introduction

For computer simulations, we propose to use SUMO (Simulation of Urban Mobility), an open source microscopic urban traffic simulator developed by the German aerospace center (DLR) in 2002 (see [13]). However, SUMO is a deterministic and uncertainty-free simulation environment. Therefore, similar to simulations conducted in [11]and [12], we propose to use equation (2) that was originally proposed in [12] to render SUMO a stochastic and uncertain simulation environment.

$$v(t + \Delta t) = v(t) + (\tau_X \times a \times \Delta t)$$
$$\tau_X = rand\{1 - \gamma, 1 - \gamma/2\} \quad (2)$$

For instance, using $\gamma = 0$ in (2) results in deterministic and accurate accelerations which simulates the driving in a deterministic and uncertainty-free environment. However in this paper, we propose to use $\gamma = 0.4$ which can reduce the accuracy of the vehicles' accelerations up to 60% and, as a result, produce a stochastic and uncertain simulation environment.

Also, it should be noted that the road network and the traffic flow used in our microscopic simulations was randomly generated using SUMO's *netgenerate* and *randomTrip*, respectively (see Fig.2).

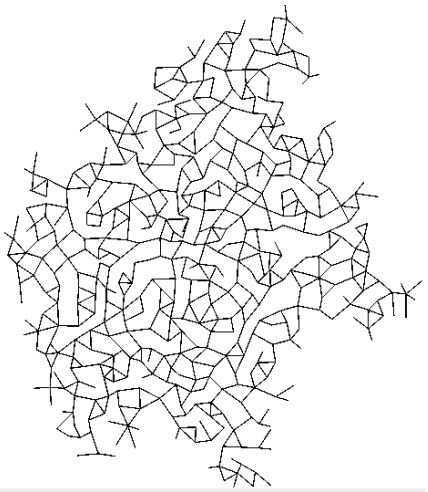


Fig. 2. A randomly generated road network and traffic flow for a more realistic simulation environment.

We present in the rest of this section some numerical results obtained from conducting several microscopic simulations in a stochastic and uncertain environment. We also provide a discussion while highlighting the key elements of learning an accurate transition model and providing collision-free and autonomous longitudinal control.

B. Learning from the HAgent's driving behavior

In this subsection, we examine the changes in the four ratio values during simulation, in particular during the exploration stage. Here in this simulation, while the $HAgent$ controls the velocity to simulate manual driving, the $DAgent$ observes its behavior to learn the environment's transition model.

It should be noted that because the outcomes of the test vehicle's accelerations are configured to be stochastic, the results presented in Fig.3 represent the average results of five simulation iterations of the same driving scenario. Here in this scenario, the test vehicle traveled the same route five times, each time the $DAgent$ starts building its transition model from scratch. This allows as to obtain a more accurate representation of the convergence of the ratio values in a stochastic and uncertain environment.

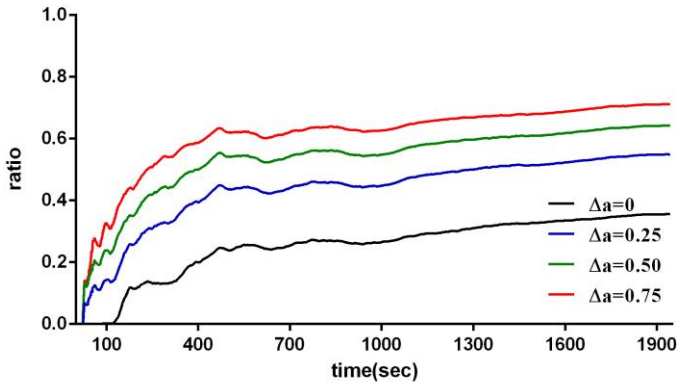


Fig. 3. The average changes in the ratio values over the simulation duration of the same driving scenario.

Based on the results presented in Fig.3, the ratio values for instance of $r_{0.75}$ are always greater than the other ratio values, and this is because in $r_{0.75}$ the $DAgent$ has more margin for error. On the other hand, the ratio values of r_0 are always lesser than the others because it does not allow any error. In fact, in r_0 we increment $n_{\Delta a=0}$ only if the decision made by the $DAgent$ is identical to the $HAgent$'s ($\Delta a_0 = 0\ m/s^2$).

While the $DAgent$ starts the exploration stage, it observes each acceleration a exercised (at $t - \Delta t$) by the $HAgent$ to include it in the transition model in addition to the vehicle state before ($v_{t-\Delta t}$) and after (v_t) exercising a. Using these three parameters as inputs, the $DAgent$ builds a transition model to acquire a representation of the inaccuracy of its accelerations in an uncertain environment. Therefore, the more inputs it perceives the more accurate this representation and, as a result, the transition model become.

Furthermore, a drop in the ratio values always indicates that the $DAgent$ has made an error, more specifically has made a decision crucially different than the $HAgent$'s optimal decision. On the contrary, the increase in the same value indicates a similarity between both agent's decisions. However, why does the autonomous $DAgent$ sometimes makes different decisions?

The only difference in the decision process of both agents is the transition model they use. The $HAgent$ uses a pre-built and accurate transition model to simulate the human driver's optimal behavior. On the other hand, the $DAgent$ uses a new transition model, a model that it builds based on its observations. Therefore, a drop in the ratio values indicates that the $DAgent$ has encountered a new situation, or due to the inaccuracy of information provided by its transition model.

On the other hand, an increase in the ratio values indicates that the $DAgent$ had plenty of interactions with its environment in a state identical to its current state. This allowed the $DAgent$ to acquire an accurate representation of the inaccuracy of its actions and, therefore, to make decisions close to the $HAgent$'s.

In the second iteration of simulating the same driving scenario, and after using the built transition model, the ratio values increased to : $r_0 = 0.51$, $r_{0.25} = 0.73$, $r_{0.50} = 0.83$ and $r_{0.75} = 0.87$. This increase in the ratio values is explained by the fact that the autonomous agent now encounters more familiar situations where it made decisions close the human driver's actions.

After the third iteration, the values increased to : $r_0 = 0.55$, $r_{0.25} = 0.82$, $r_{0.50} = 0.90$ and $r_{0.75} = 0.93$. This indicates that the decisions made by the autonomous agent have become more accurate, more specifically have got more close to the human agent's. For instance, during the third iteration and in almost 55% of the times, the $DAgent$ made decisions identical to the $HAgent$'s actions.

C. Handling control uncertainty

After three simulation iterations, and even after reaching a ratio of 55%, the $DAgent$ still causes collisions on some occasions during the autonomous driving. This can be due to encountering other new situations where the agent can not accurately estimate the outcome of its actions, or due to the insufficiency of its interactions with the uncertain environment.

For instance, in a new simulation scenario where basically the $DAgent$ autonomously drives the vehicle in a different traffic configuration using the built transition model, the agent caused a collision with another vehicle. We present in the figures below the numerical results of the simulation, as well as a comparison of both the $DAgent$'s behavior and the $HAgent$'s optimal behavior during the last forty steps before the collision.

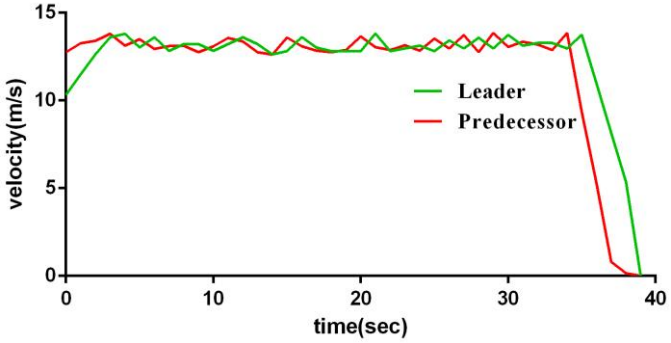


Fig. 4. The following vehicle's velocity and the preceding vehicle's

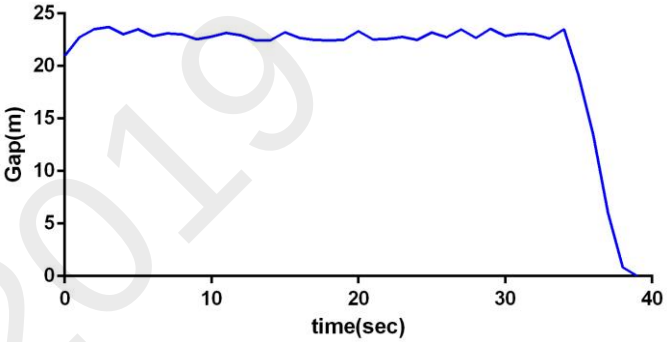


Fig. 5. The gap between the two vehicles.

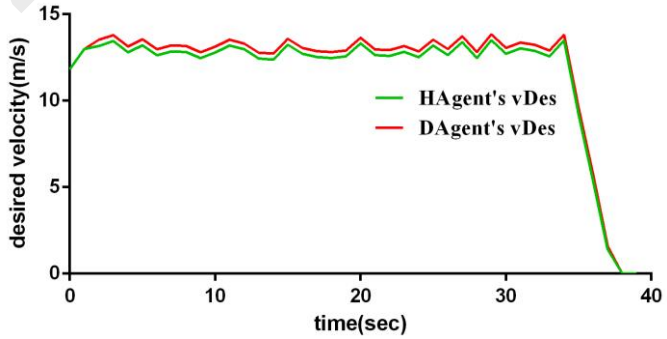


Fig. 6. The safe velocity defined by each agent.

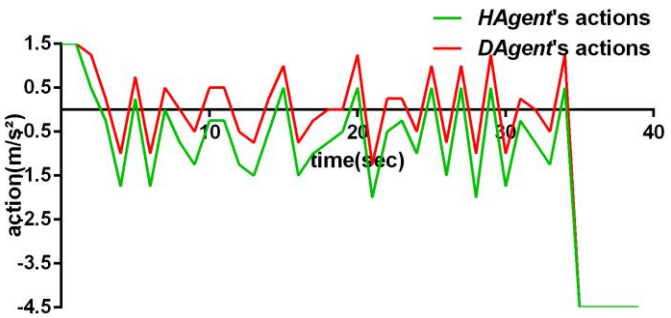


Fig. 7. The $DAgent$'s behavior compared to $HAgent$'s optimal behavior.

Here, based on the results presented above and using as a reference the $HAgent$'s behavior which we consider to be optimal, we examine the behavior of the $DAgent$ to identify the cause of collision.

During simulation, both agents define a desired velocity ($vDes$) that should allow them to maintain collision-free longitudinal control. However, during the last forty seconds before the collision, the $DAgent$ defined a desired velocity that was always greater than the $HAgent$'s (see Fig.6). This made the $DAgent$ think that it is still far from the collision and it can drive faster to shorten the gap to the preceding vehicle. As a result, the $DAgent$ exercised acceleration values greater than those chosen by the $HAgent$'s which eventually caused the collision (see Fig.7). However, what is the cause behind the miss estimation of the safe and desired velocity $vDes$?

Both the $HAgent$ and the $DAgent$ use the same equation and the same reward function to control the velocity of the vehicle. But, they both use different transition models, the $HAgent$ uses a pre-built and accurate transition model to simulate an optimal driving behavior, whereas the $DAgent$ uses a transition model built from its previous interactions with the environment.

Therefore, we explain the cause of the miss estimation of the desired velocity as well as the cause of the collision by the miss estimation of the accelerations' outcomes using an inaccurate transition model. In other words, the autonomous agent didn't interact enough in the new traffic configuration to accurately estimate the outcome of its actions and, even more importantly, to accurately define the safe velocity.

Also, it should be noted that after the end of the simulation in the new traffic configuration, the ratio values dropped to: $r_0 = 0.36$, $r_{0.25} = 0.42$, $r_{0.50} = 0.52$ and $r_{0.75} = 0.78$. Furthermore, after three more iterations, the $DAgent$ finally achieved collision-free and autonomous longitudinal control in the new traffic configuration under control uncertainty.

IV. CONCLUSION

Providing a fully autonomous and collision-free velocity control require the learning of both the human driver's reward function and the environment's transition model. However, we focused in this paper on learning the transition model and examining the effect of its inaccuracy on the performance of the autonomous longitudinal control system.

First, we note that the performance of the $DAgent$ can be significantly reduced by newly encountered situations, situations where this agent has never interacted or didn't interact enough with the environment. In such situations, the estimation of inaccuracy of its actions can be inaccurate and, as a result, the autonomous system could cause several collisions due.

On the other hand, in the case of encountering familiar situations where the vehicle had plenty of interactions with the uncertain environment, the $DAgent$ maintained an autonomous and collision-free longitudinal control. Also, using the ratio values reflect only the agent's performance in the past encountered states, and does not describe the system's ability to produce a collision-free driving in the future.

To this end, building an accurate transition model is crucial to providing an autonomous and collision-free longitudinal control. This requires the autonomous system to explore all the possible outcomes of exercising any given action in any given state, which can be achieved after a long period of observation and learning.

Finally, it should be noted that the system proposed in this paper focuses only on providing collision-free and autonomous longitudinal control. However, we propose in the future to:

- Learn the human driver's reward function in order to provide a more accurate, realistic and efficient longitudinal control under uncertainty,
- Propose another approach to evaluate the $DAgent$'s ability to handle control uncertainty,
- Incorporate the automation of the lateral control towards fully automating the driving task under control uncertainty,

REFERENCES

[1] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards, Artificial intelligence: a modern approach vol. 3: Prentice hall Upper Saddle River, 2010.

[2] A. González-Sieira, M. Mucientes, and A. Bugarín, "A state lattice approach for motion planning under control and sensor uncertainty," in ROBOT2013: First Iberian robotics conference, 2014, pp. 247-260.

[3] S. Brechtel, T. Gindele, and R. Dillmann, "Probabilistic decision-making under uncertainty for autonomous driving using continuous POMDPs," in Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on, 2014, June, pp. 392-399.

[4] Y. W. Seo and C. Urmson, "A perception mechanism for supporting autonomous intersection handling in urban driving," in Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, 2008, pp. 1830-1835.

[5] J. Wei, J. M. Dolan, J. M. Snider, and B. Litkouhi, "A point-based mdp for robust single-lane autonomous driving behavior under uncertainties," in ICRA, 2011, pp. 2586-2592.

[6] A. Vahidi and A. Eskandarian, "Research advances in intelligent collision avoidance and adaptive cruise control," IEEE transactions on intelligent transportation systems, vol. 4, pp. 143-153, 2003.

[7] J. R. Cho, J. H. Choi, W. S. Yoo, G. J. Kim, and J. S. Woo, "Estimation of dry road braking distance considering frictional energy of patterned tires," Finite Elements in Analysis and Design, vol. 42, pp. 1248-1257, 2006.

[8] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, et al., "Stanley: The robot that won the DARPA Grand Challenge," Journal of field Robotics, vol. 23, pp. 661-692, 2006.

[9] T. Helldin, G. Falkman, M. Riveiro, and S. Davidsson, "Presenting system uncertainty in automotive uis for supporting trust calibration in autonomous driving," in Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications, 2013, pp. 210-217.

[10] A. Lahlouhi, "Modélisation multi-agent du Processus logiciel," Thèse de doctorat, Département d'Informatique, Faculté des Sciences de l'Ingénieur, Université Mentouri, Algérie, 2006.

[11] O. Messaoudi, "An optimal velocity robust car-following model with consideration of control uncertainty," in 2018 International Conference on Applied Smart Systems (ICASS), 2018, pp. 1-8.

[12] O. Messaoudi and A. Lahlouhi, "An agent-based inter-vehicle cooperative robust car-following model for longitudinal control under uncertainty," International Journal of Computer Applications in Technology, vol. 58, pp. 150-164, 2018.

[13] D. Krajzewicz, J. Erdmann, M. Behrisch, and L. Bieker, "Recent development and applications of SUMO–simulation of urban mobility," International Journal On Advances in Systems and Measurements, vol. 5, 2012.

Some remarks about optimal control of PDEs with missing data

Abdelhak Hafdallah
Department of mathematics, University of Tebessa.
University of Tebessa. Rue de Constantine12002
Algeria
hafdallahmath@gmail.com

Mereim Louafi
Department of mathematics, University of Tebessa.
University of Tebessa. Rue de Constantine12002
Algeria
mirou420@yahoo .fr

Mouna Abdelli
Department of mathematics, University of Tebessa.
University of Tebessa. Rue de Constantine12002
Algeria
mouna9393@gmail.com

Abstract— **The aim of this paper leads to introduce a new and equivalent definition for the no-regret control introduced by J. L. Lions to study optimal control of equations with incomplete data. As a result, on the contrary of the original definition, the new definition is beneficial because it leads to a simpler optimality system. We apply the new definition to control some PDEs with missing data.**

Keywords—**optimal control, no-regret control, incomplete data, PDEs**.

I. INTRODUCTION

When modeling some natural phenomena we cannot have access to all data, due to the inability to measure her or because of big errors in measures, these reasons lead to get some incomplete data models or partially incomplete data models. The optimal control study of that kind of problems has received much attention in the last three decades starting by J. L. Lions in [7, 10] where he introduced the notions of no-regret control and low-regret, the no-regret concept has originated in statistics by Savage [11].

Later, few studies have been published after that by Lions himself like [9], then many authors are attracted to apply the no-regret control to study different optimal control problems with missing data, as in nonlinear systems with incomplete data [12], distributed systems [13], also in population dynamics with incomplete data [3]. Recently, works studied fractional diffusion, diffusion equations both with some missing data its model in [5, 6] and [14] respectively. Moreover, in [2] authors coupled the concepts of no-regret control and averaged control (see [4]) to study a electromagnetic equation with two types of missing data, and they also studied a coupled system from thermoelasticity with missing initial conditions in [1].

The present paper presents a new and equivalent definition for the no- regret, we remember that the original one has been given by J. L. Lions in [7] to solve optimal control problems with missing data. Actually, the novel definition of no-regret control leads to a simpler optimality system i.e., a reduced structure of optimality system which is expected to simplify the old characterization of no-regret control by minimizing the number of state equations in optimality systems.

Consequently, the new definition will make the numerical analysis of optimal control problems with incomplete data easier to do, this is in the sense of minimizing the number of operations in numerical treatments.

The outline of this paper is as follows: in the next section we present the original definition of no-regret related to some fixed control, in the third section devoted to introduce the new equivalent definition of no-regret control, in the fourth we present a characterization for the no-regret control based on the new definition, in the last section we give a characterization for low regret control and no-regret control, in the last section we shall apply the main results to control some PDEs with incomplete data, and finally we end the paper with a conclusion.

II. PRELIMINARIES AND DEFINITIONS

Consider the following linear state equation described by :

$$Ay(v,g) = Bv + \beta g, \qquad (1)$$

where AϵL(V,V') is a partial differential operator isomorphic from a Hilbert space V to its dual V', B ϵ L (U,V') is a control operator, U is a Hilbert space of controls, βϵL(G,V') where G is a Hilbert space of missing data, v ϵ U is the control function and g ϵ G is a missing data.

Suppose that (1) is well posed in V and denote by y (v, g) her unique solution depending on the control v and on the missing data g.

Associate to (1) the objective quadratic function of the form (see [8])

$$J_0(v,g) = \|Cy(v,g) - z_d\|_H^2 + N\|v\|_U^2 \qquad (2)$$

where CϵL (V,H), H is a Hilbert space, z_d is a desired state in H and N > 0.

Our goal is to characterize the optimal control of (1) subject to the cost function (2) whatever the value of the missing data g, i.e., we are looking for an optimal control independently of g. In other words, we are looking to solve

$$\inf_{v\epsilon U} J_0(v,g) \quad for\ evry\ g\epsilon G$$

This definition doesn't make any sense when $G \neq \{0\}$. One thinks to look for the following min-max problem

$$\inf_{v\epsilon U}\left(\sup_{g\epsilon G} J_0(v,g)\right) \qquad (3)$$

but we can get $\sup_{g\epsilon G} J_0(v,g) = +\infty$,in this case (3) has no sense.

The last difficulty leads J.L.Lions to think about looking for controls such that

$$J_0(v,g) \leq J_0(u_0,g) \text{ for evry } g \in G \qquad (4)$$

where $u_0 \in U$. Note that in (4), we have chosen the controls v that do better than u_0, note also that the optimal control belongs to this set of controls.

Now, let's call back the original definition of no-regret control.

Definition 1 We say that $u \in U$ is a no-regret control related to $u_0 \in U$ for (1)(2) if u is a solution of

$$\inf_{v\in U}\left(\sup_{g\in G}\left(J_0(v,g) - J_0(u_0,g)\right)\right.$$

In our recent work, we suggest an alternative way of thinking to get a new and equivalent definition of no-regret as follows:

By of linearity in (1), we have

$$J_0(v,g) = \|Cy(v,0) + Cy(0,g) - z_d\|_H^2 + N\|\text{by}$$

looking at the cost function J_0 as a function of v and y(0,g) instead of v and g, we can write:

$$J_0(v,g) = \|Cy(v,0) + Cy(0,g) - z_d\|_H^2 + N\|v\|_U^2$$

where the function J is given by

$$J_0\big(v,y(0,g)\big) - J_0(v,g) = \|Cy(v,g) - z_d\|_H^2 + N\|v\|_U^2$$

Let $Y=\{y(0,g),\ g\in G\}$ be a subspace of V, the key idea is to remark that

$$\sup_{g\in G} J_0\big(v,y(0,g)\big) = \sup_{y(0,g)\in Y} J_0\big(v,y(0,g)\big)$$

this allows us to rewrite the min-max problem (3) as follows

$$\inf_{v\in U}\left(\sup_{y(0,g)\in Y} J\big(v,y(0,g)\big)\right) \qquad (5)$$

III. NO-REGRET CONTROL AND LOW-REGRET CONTROL NEW DEFINITIONS

Now, let's give a new and equivalent definition of no-regret control based on (5) as follows:

Definition 2 Let u_0 be a fixed control in U. We say that $u \in U$ is a no-regret control for (1) (2) related to u_0 if u is a solution of

$$\inf_{v\in U}\left(\sup_{y(0,g)\in Y}\left(J\big(v,y(0,g)\big) - J\big(u_0,y(0,g)\big)\right)\right) \qquad (6)$$

Remark 3 Note that definition 2 takes into account the new cost function form J(v,y(0,g)) instead of its old form $J_0(v,g)$, i.e., the missing state y(0,g) will play the role of the missing data g.

Then, we'll try to rewrite the last quantity under inf-sup in new form to separate the roles of v and y(0,g) by the following

Lemma 4 For every $(v,g)\in U\times G$, we have

$$J\big(v,y(0,g)\big) - J\big(u_0,y(0,g)\big) = J(v,0) - J(u_0,0) + 2(C^*Cy(v-u_0,0),y(0,g))_{V'} \qquad (7)$$

Proof Because J(v,y(0,g))=J_0(v,g) 0and y(0,0)=0 and by a simple calculation we get (7).

Relax problem (6) by making a quadratic perturbation on the missing state, i.e., looking only for controls such that

$$J\big(v,y(0,g)\big) - J\big(u_0,y(0,g)\big) \leq \gamma\|y(0,g)\|_V^2$$

for evry $g\in G$ with $\gamma > 0$

Let's give an alternative definition of low-regret control (for the original one see [7]) as follows :

Definition 5 We say that $u_\gamma \in U$ is a low-regret control related to $u_0 \in U$ for (1)(2) if u_γ is a solution of

$$\inf_{v\in U}\left(\sup_{y(0,g)\in Y}\left(J\big(v,y(0,g)\big) - J\big(u_0,y(0,g)\big) - \gamma\|y(0,g)\|_V^2\right)\right) \qquad (8)$$

In that case, the above definition could be written in a different form as follows:

By using (6) we get

$$\sup_{y(0,g)\in Y}\left(J\big(v,y(0,g)\big) - J\big(u_0,y(0,g)\big) - \gamma\|y(0,g)\|_V^2\right) = J(v,0) - J(u_0,0) + \sup_{y(0,g)\in Y}\left(2\big(C^*Cy(v-u_0,0),y(0,g)\big)_V - \gamma\|y(0,g)\|_V^2\right) \leq J(v,0) - J(u_0,0) + \sup_{y\in V}(2(C^*Cy(v-u_0,0),y)_V - \gamma\|y\|_V^2) = J(v,0) - J(u_0,0) + \frac{1}{\gamma}\|C^*Cy(v-u_0,0)\|_{V'}^2$$

Identify the Hilbert space V to its dual space V' to obtain a new form for our optimal control problem,

$$\inf_{v\in U} J^\gamma(v) \text{ where } J^\gamma(v) = J(v,0) - J(u_0,0) + \frac{1}{\gamma}\|C^*Cy(v-u_0,0)\|_V^2 \qquad (9)$$

Finally, we have gotten a classical optimal control problem that depends only on the control function v, then, we are able to apply the classical theory to obtain some optimality system characterizing the solution of (9).

IV. LOW-REGRET CONTROL AND NO-REGRET CONTROL CHARACTERIZATIONS (NEW OPTIMALITY SYSTEMS)

First, Let's give an existence-uniqueness result given in the following

Proposition 6 The optimal control problem (9) has a unique solution u_γ for every $u_0 \in U$.

Proof We have $J_\gamma(v) \geqslant -J(0,0)$=constant for every $v\in U$ then $d_\gamma = \inf J_\gamma(v)$ exists. Let $v_n = v_n(\gamma)$ be a minimizing sequence with $J_\gamma(v_n) \to d_\gamma$ then

$$-J(0,u_0) \leq J(v_n,0) - J(u_0,0) \frac{1}{\gamma}\|C^*Cy(v-u_0,0)\|_{V'}^2 \leq d^\gamma + 1$$

from this we deduce $\|v_n\|_U \leqslant C_\gamma$ independent of n. Then, there exists $u_\gamma \in U$ such $v_n \to u_\gamma$ weakly in U. Also, by continuity w.r.t data $y(v_n,0) \to y(u_\gamma,0)$ in V. Moreover, from strict convexity of J_γ we deduce that u_γ is unique.

It remains to prove that the sequence u_γ converges to u the no-regret control related to u_0 when $\gamma \to 0$.

Theorem 7 The sequence u_γ solution to (9) converges to the no-regret control u related to u_0 weakly in U when $\gamma \to 0$.

Proof We know that u_γ solves (9), then for every v∈U we have

$$J(u_\gamma,0)-J(u_0,0)+\frac{1}{\gamma}\left\|C^*Cy(u_\gamma-u_0,0)\right\|_{V'}^2 \le J(v,0)-J(u_0,0)\frac{1}{\gamma}\|C^*Cy(v-u_0,0)\|_{V'}^2,$$

choose v=u_0 to find

$$\left\|Cy(u_\gamma-u_0,g)-z_d\right\|_H^2+N\|u_\gamma\|_U^2+\frac{1}{\gamma}\left\|C^*Cy(u_\gamma-u_0,0)\right\|_{V'}^2 \le J(u_0,0)=constant$$

which implies the following bounds for some c>0

$$\left\|Cy(u_\gamma-u_0,g)\right\|_H \le c,\ \|u_\gamma\|_U \le c,\ \left\|C^*Cy(u_\gamma-u_0,0)\right\|_{V'} \le c\sqrt{\gamma} \quad (10)$$

which from we deduce that the sequence u_γ is bounded in U then we can extract a subsequence still be denoted u_γ that converges weakly to u∈U.

It remains to prove that u is a no-regret control related to u_0. It's clear that for every v∈U

$$J(v,g)-J(u_0,g)-\gamma\|y(0,g)\|_V^2 \le J(v,g)-J(u_0,g)\ for\ evry\ g\in G$$

then

$$J(u_\gamma,g)-J(u_0,g)-\gamma\|y(0,g)\|_V^2 \le \sup_{y(0,g)\in Y}(J(v,g)-J(u_0,g))$$

Make $\gamma\to 0$ to find

$$\sup_{y(0,g)\in Y}(J(u,g)-J(u_0,g)) \le \sup_{y(0,g)\in Y}(J(v,g)-J(u_0,g))$$

the last inequality is equivalent to saying that

$$\sup_{y(0,g)\in Y}(J(u,g)-J(u_0,g))=\inf_{v\in U}(\sup_{y(0,g)\in Y}(J(v,g)-J(u_0,g$$

i.e., u is a no-regret control related to u_0.

The following theorem will give an optimality system for uγ solution to (9).

Theorem 8 The low-regret control uγ related to u_0∈U solution to (9) is characterized by

$$\begin{cases} Ay_\gamma = Bu_\gamma \\ A^*\mathcal{C}_\gamma = C^*Cy_\gamma - z_d + \frac{1}{\gamma}(C^*C)^2y(u_\gamma-u_0,0), \\ B^*\mathcal{C}_\gamma + Nu_\gamma = 0\ \ in\ U \end{cases} \quad (11)$$

where yγ=y(uγ,0).

Proof A first order optimality condition gives for every w∈U

$$(Cy_\gamma-z_d,Cy(w,0))_H+N(u_\gamma,w)_U+\frac{1}{\gamma}(C^*Cy(u_\gamma-u_0,0),C^*Cy(w,0))_V \ge 0 \quad (12)$$

or

$$(C^*Cy_\gamma-z_d+\frac{1}{\gamma}(C^*C)^2y(u_\gamma-u_0,0),y(w,0))_V+N(u_\gamma,w)_U \ge 0$$

Introduce the adjoint state $\mathcal{C}_\gamma$ ∈V by

$$A^*\mathcal{C}_\gamma = C^*Cy_\gamma - z_d+\frac{1}{\gamma}(C^*C)^2y(u_\gamma-u_0,0)$$

then, (12) is equivalent to the following inequality

$$(B^*\mathcal{C}_\gamma+Nu_\gamma,w)_V \ge 0$$

don't forget that U is a vector space, hence

$$B^*\zeta_\gamma+Nu_\gamma=0.$$

Finally, we could give an optimality system characterizing u the no-regret control related to u_0.

Theorem 9 The no-regret control u related to u_0 is characterized by the following optimality system

$$\begin{cases} Ay = Bu \\ A^*\zeta = C^*Cy(u,0)-z_d+\lambda, \\ B^*\zeta+Nu=0\ \ in\ U \end{cases} \quad (13$$

where λ ∈V.

Proof From (10) we know that uγ⇀u weakly in U, and by continuity of B from U into V' we conclude that Buγ⇀Bu in V'. Also, from optimality system (11) Ayγ is bounded in V′ then weakly convergent to Ay (because A is an isomorphism), then by passing to limit we get Ay=Bu. By the same way, we prove also that

$$C^*Cy_\gamma+(1/\gamma)(C^*C)^2y(u_\gamma-u_0,0) \rightharpoonup C^*Cy+\lambda\ \ weakly\ in\ V'$$

We deduce also that A*$\mathcal{C}$ γ is bounded in V′ which implies the boundedness of $\mathcal{C}$ γ in V therefore

$$A^*\mathcal{C}_\gamma \rightharpoonup A^*\mathcal{C}_\gamma$$

weakly in V′, by limit uniqueness we get

$$A^*\mathcal{C}=C^*Cy(u,0)-z_d+\lambda.$$

For the last equality, by the boundness of uγ, $\mathcal{C}$ γ and by passing to limit we find:

$$B^*\mathcal{C}+Nu=0\ \ in\ U$$

Generally speaking, advantages in terms of the new definition of no-regret control far outweigh the advantages with regard to the original definition of no-regret control. The key benefits is the simpler form of optimality system characterizing the no-regret control where comparing with the old definitions in (see [7,13]), this makes numerical treatments easier to do.

V. Application to some optimal control problems with incomplete data

In this section, we try to characterize the no-regret control related to u_0=0 by using the equivalent definition 2, for different kinds of optimal control problems with incomplete data. For every case, we'll find an optimality system characterizing the no-regret control.

Example 10 Let's consider the following elliptic equation with a distributed control action and a missing Newmann boundary condition

$$\begin{cases} -\Delta y+y=v\ in\ \Omega \\ \frac{\partial y}{\partial v}=g\ on\ \Gamma \end{cases} \quad (14$$

where Ω is bounded set in $\mathbb{R}^n$ with a regular boundary Γ, v∈U=$L^2(\Omega)$ and g∈G=$L^2(\Gamma)$. It's well known that (14) has a unique solution y(v,g) is unique in $H^{3/2}(\Omega)$.Associate to (14) the following cost function

$$J(v,g)=|y(v,g)-z_d|_{L^2(\Gamma)}^2+N\|v\|_{L^2(\Omega)}^2. \quad (15)$$

Note that

$$J(v,g) - J(0,g) = J(v,0) - J(0,0) + 2\int_{\Gamma} y(v,0)y(0,g)\,d\Gamma,$$

where dΓ denote the Lebesgue measure on the boundary Γ.
The low-regret control is the solution of the following optimal control problem

$$inf_{v\in L^2(\Omega)} J^{\gamma}(v) \text{ with } J^{\gamma}(v) = J(v,0) - J(0,0) + \frac{1}{\gamma}|y(v,0)|^2_{L^2(\Gamma)}. \quad (16$$

Start by the following theorem :

Theorem 11 The low-regret control uγ solution to (16) is unique and characterized by the following optimality system

$$\begin{cases} -\Delta y_{\gamma} + y_{\gamma} = u_{\gamma}; -\Delta \zeta_{\gamma} + \zeta_{\gamma} = 0 \text{ in } \Omega \\ \dfrac{\partial y_{\gamma}}{\partial v} = 0; \dfrac{\partial \zeta_{\gamma}}{\partial v} = y_{\gamma} - z_d + \dfrac{1}{\gamma} y_{\gamma} \text{ on } \Gamma \\ \zeta_{\gamma} + Nu_{\gamma} = 0 \text{ n } L^2(\Omega) \end{cases}$$

Proof A first order optimality condition for (15) gives

$$(y(u_{\gamma},0) + \frac{1}{\gamma} y(u_{\gamma},0), y(w,0))_{L^2(\Gamma)} + N(u_{\gamma},w)_{L^2(\Omega)} \geq 0, \forall w \in L^2(\Omega).$$

Introduce an adjoint state ζγ solution of

$$\begin{cases} -\Delta \zeta_{\gamma} + \zeta_{\gamma} = 0 \text{ in } \Omega \\ \dfrac{\partial \zeta_{\gamma}}{\partial v} = y_{\gamma} - z_d + \dfrac{1}{\gamma} y_{\gamma} \text{ on } \Gamma \end{cases}$$

then, use Green formula to get

$$\zeta_{\gamma} + Nu_{\gamma}, w)_{L^2(\Omega)} \geq 0.$$

To obtain no-regret control optimality system, adapt the proof of theorem 8 to find the following result:

Theorem 12 The no-regret control u for (14)(15) is unique and characterized by

$$\begin{cases} -\Delta y + y = u; -\Delta \zeta + \zeta = 0 \text{in } \Omega \\ \dfrac{\partial y}{\partial v} = 0; \dfrac{\partial \zeta}{\partial v} = y - z_d + \lambda \quad \text{on } \Gamma \\ \zeta + Nu = 0 \text{ in } L^2(\Omega) \end{cases}$$

where $\lambda \in L^2(\Gamma)$.

Example 13 It's a fourth order parabolic equation with a distributed control and a missed initial condition, given by

$$\begin{cases} y + \Delta(a(x,t)\Delta y) = v \text{ in } Q \\ y = 0, \dfrac{\partial y}{\partial v} = 0 \text{ on } \Sigma \\ y(0) = g \text{ in } \Omega \end{cases} \quad (17$$

where Ω is an open set in $\mathbb{R}^n$ with a smooth boundary

$$\Gamma, t \in [0;T], T > 0, Q = \Omega \times]0,T[\text{ and } \Sigma = \Gamma \times]0,T[. \text{With } a \in L^{\infty}(Q), a \geq \alpha > 0$$

almost everywhere, v is a control function in U=L²(Q) and g is a missing initial data in G=L²(Ω). The equation (17) has a unique solution in $L^2(0,T;H_0^2(\Omega))$ (see [8]). Consider the following cost function

$$J(v,g) = \|y(v,g) - z_{\{d\}}\|^2_{L^2(Q)} + N\|v\|^2_{L^2(Q)}. \quad (18$$

The low-regret control is a solution of

$$inf_{v\in L^2(Q)} J^{\gamma}(v) \text{ with } J^{\gamma}(v) = J(v,0) - J(0,0) + \frac{1}{\gamma}\|y(v,0)\|^2_{L^2(Q)}. \quad (19$$

Theorem 14 The low-regret control u_{γ} solution to (19) is unique and it's characterized by the following optimality system

$$\begin{cases} y_{\gamma}' + \Delta(a(x,t)\Delta y_{\gamma}) = u_{\gamma} \text{on } \Sigma \\ -\zeta_{\gamma}' + \Delta(a(x,t)\Delta \zeta_{\gamma}) = y_{\gamma} - z_d + \dfrac{1}{\gamma} y_{\gamma} \text{ in } Q \\ y_{\gamma} = 0, \left(\dfrac{\partial y_{\gamma}}{\partial v}\right) = 0 \text{ on } \Sigma \\ \zeta_{\gamma} = 0, \dfrac{\partial \zeta_{\gamma}}{\partial v} = 0 \text{ on } \Sigma \\ y_{\gamma}(0) = 0; \zeta_{\gamma}(T) = 0 \text{ in } \Omega \end{cases}$$

Where $y_{\gamma} = y(u_{\gamma},0)$, with $\zeta_{\gamma} + Nu_{\gamma} = 0 \text{ in } L^2(Q)$.

Proof Again, a first order optimality condition gives for every w∈L² (Q)

$$(y(u_{\gamma},0) - z_d + \frac{1}{\gamma}\left(y(u_{\gamma},0), y(v - u_{\gamma},0)\right)_{L^2(Q)} + N(u_{\gamma}, v - u_{\gamma})_{L^2(L^2(Q))} \geq 0$$

Introduce an adjoint state ζγ} given by

$$\begin{cases} -\zeta_{\gamma}' + \Delta(a(x,t)\Delta \zeta_{\gamma}) = y_{\gamma} - z_d + \dfrac{1}{\gamma} y_{\gamma} \text{in } Q \\ \zeta_{\gamma} = 0, \dfrac{\partial \zeta_{\gamma}}{\partial v} = 0 \text{on } \Sigma \\ \zeta_{\gamma}(T) = 0 \text{in } \Omega \end{cases}$$

then

$$\zeta_{\gamma} + Nu_{\gamma} = 0 \text{ in } L^2(Q).$$

Theorem 15 The no-regret control u for (17)(18) is unique and it's characterized by the following optimality system

$$\begin{cases} y' + \Delta(a(x,t)\Delta y) = u \text{in } Q \\ -\zeta' + \Delta(a(x,t)\Delta \zeta) = y(u,0) - z_d + \lambda \text{in } Q \\ y = 0, \dfrac{\partial y}{\partial v} = 0; \zeta = 0, \dfrac{\partial \zeta}{\partial v} = 0 \text{on } \Sigma \\ y(0) = 0; \zeta(T) = 0 \text{ in } \Omega \end{cases}$$

where $\zeta + Nu = 0 \text{ in } L^2(Q)$.

Example 16 At last, let's take a hyperbolic example: it's a wave equation with a boundary control action and a missing source

$$\begin{cases} y'' - \Delta y = g \\ \left(\dfrac{\partial y_{\gamma}}{\partial v}\right) = 0 \\ y(0) = 0 \text{ ; } y'(0) = 0 \end{cases} \quad (20$$

where Ω is an open in $\mathbb{R}^n$ with a smooth boundary Γ, t∈[0;T], T>0, Q=Ω×]0;T[, Σ=Γ×]0;T[, v is a boundary control in U=L²(Σ) and g∈G=L²(Q) is an unknown function. By

transposition method, there is a unique solution $y \in L^2(Q)$ (see [8]. Associate to (20) the cost function

$$J(v,g) = \|y(v,g) - z_d\|^2_{L^2(Q)} + N\|v\|^2_{L^2(\Sigma)}, \quad (21$$

the low-regret control is a solution of

$$inf_{v \in L^2(\Sigma)} J^\gamma(v) \; with J^\gamma(v) = J(v,0) - J(0,0) + \frac{1}{\gamma}\|y(v,0)\|^2_{L^2(Q)}. \quad (22$$

Theorem 17 The low-regret control uγ solution to (22) is characterized by the following optimality system

$$\begin{cases} y_\gamma'' - \Delta\, y_\gamma = 0; \text{in } Q \\ \zeta_\gamma'' - \Delta\, \zeta_\gamma = y_\gamma - z_d + \frac{1}{\gamma} y_\gamma \text{in } Q \\ \frac{\partial y_\gamma}{\partial v} = u_\gamma; \frac{\partial\, \zeta_\gamma}{\partial v} = 0 \text{on } \Sigma, \\ y_\gamma(0) = 0, y_\gamma'(0) = 0;\ \zeta_\gamma(T) = 0,\ \zeta_\gamma'(T) = 0 \text{in } \Omega \end{cases}$$

with $\zeta_\gamma + Nu_\gamma = 0 \; in \; L^2(\Sigma)$.

Proof Another time, a first order optimality condition gives for every $w \in L^2(\Sigma)$

$$\left(y_\gamma - z_d + \frac{1}{\gamma} y_\gamma, y(w,0)\right)_{L^2(Q)} + N\left(u_\gamma, w\right)_{L^2(\Sigma)} 0,$$

define an adjoint state ζγ by

$$\begin{cases} \zeta_\gamma'' - \Delta \zeta_\gamma = y_\gamma - z_d + \frac{1}{\gamma} y_\gamma \text{in } Q \\ \frac{\partial \zeta_\gamma}{\partial v} = 0 \text{ on } \Sigma \\ \zeta_\gamma(T) = 0, \zeta_\gamma'(T) = 0 \text{in } \Omega \end{cases}$$

to find

$$\zeta_\gamma + Nu_\gamma = 0 \quad in \; L^2(\Sigma).$$

Theorem 18 The no-regret control u for (21)(22) is characterized by

$$\begin{cases} y'' - \Delta\, y = 0; \zeta'' - \Delta\zeta = y - z_d + \lambda \text{in } Q \\ \frac{\partial y}{\partial v} = u; \frac{\partial \zeta}{\partial v} = 0 \text{on } \Sigma \\ y(0) = 0, y'(0) = 0;\ \xi(T) = 0, \zeta'(T) = 0 \text{ in } \Omega \end{cases}$$

where $\lambda \in L^2(Q)$, and $\zeta + Nu = 0 \; in \; L^2(\Sigma)$.

Remark 19 Moreover, the new equivalent definition of no-regret control could be applied to restudy many treated optimal control problems with missing data like (7), the new definition will simplify its study.

Conclusion

To sum up, the equivalent definition of no-regret control lead us to optimality systems which have a similar form to the classical systems [8], in contrast to the original definition [7] which have more complicated characterization, this difference is very beneficial in point of view of numerical analysis.

REFERENCE

[1] A. Hafdallah, A. Ayadi, Optimal Control of a thermoelastic body with missing initial conditions; International Journal of Control, DOI: 10.1080/00207179.2018.1519258 (2018).

[2] A. Hafdallah, A. Ayadi, Optimal control of electromagnetic wave displacement with an unknown velocity of propagation, International Journal of Control, DOI: 10.1080/00207179.2018.14581 (2018).

[3] B. Jacob, A. Omrane, Optimal control for age-structured population dynamics of incomplete data. J. Math. Anal. Appl. 370 (2010) 42--48.

[4] E. Zuazua, Averaged control. Automatica, 50(12), pp.3077-3087 (2014).

[5] D. Baleanu, C. Joseph and G. Mophou, Low-regret control for a fractional wave equation with incomplete data. Advances in Difference Equations, doi:10.1186/s13662-016-0970-8, 2016.

[6] G. Mophou, Optimal for fractional diffusion equations with incomplete data. J. Optim.Theory Appl.(2015). doi:10.1007/s10957-015-0817-6.

[7] J. L. Lions, Contrôle à moindres regrets des systèmes distribués. C. R. Acad. Sci. Paris Ser. I Math., Vol. 315, pp 1253-1257(1992).

[8] J.L. Lions, Contrôle optimal de systèmes gouvernés par des équations aux dérivées partielles. Dunod, Paris, 1968.

[9] J. L. Lions, Duality arguments for Multi Agents Least-Regret Control. Institut de Frace, Paris (1999).

[10] J. L. Lions, No-regret and low regrt control, Envirenment, Economics and Their Mathematical Models, Masson, Paris, 1994.

[11] L.J. Savage, The foundations of statistics, Dover, 1972.

[12] O. Nakoulima, A. Omrane and J. Velin, No-regret control for nonlinear distributed systems with incomplete data. J. Math. Pures Appl. 81, 1161-1189 (2002).

[13] O. Nakoulima, A. Omrane, J.Velin. On the Pareto control and no-regret control for distributed systems with incomplete data. SIAM J. CONTROL OPTIM. Vol. 42, No. 4, pp. 1167--1184 (2003).

[14] S. Mahoui, M.S. Moulay and A. Omrane. Pointwise Optimal Control for Diffusion Problems of Missing Data. Mediterranean Journal of Mathematics,14(3),p.120(2017).

Energy-Consumption-Aware Modelling and Performance Evaluation for EH-WSNs

OUKAS Nourredine[1,2] , BOULIF Menouar[3]

May 20, 2019

[1] Department of Computer Science, Normal Superior School, Kouba, Algiers, Algeria.
Email: oukas@ens-kouba.dz
[2] LIMOSE, (Computer Science/ Physics/ Mathematics) Departments.
M'Hamed Bougara University of Boumerdès, Independence Avenue, 35000, Boumerdès. Algeria.

[3] Department of Computer Science, Faculty of Sciences, M'Hamed Bougara University of Boumerdès
Email: m.boulif@univ-boumerdes.dz

Abstract

This paper proposes a Petri net modelling for the wireless sensor networks with energy harvesting capabilities. Each sensor has got an energy recovery system relying on a rechargeable battery. The energy stored in the battery is represented by levels. We use GSPN to model the effect of the communication traffic on the battery energy level. By using TimeNet tool, we show that the proposed model can simulate the network and evaluate its performance in order to get the best parameter values to ensure continuity of service for the longest possible period of time.

Key words: wireless sensor network, energy harvesting, rechargeable battery, GSPN

1 Introduction

A sensor is a small device that can collect data from the environment and send it through a wireless communication network. The main problem of these sensors is their short lifetime triggered by the limited capacity of their batteries [1, 2]. A relatively new technique consists of using rechargeable batteries to get energy from the environment (sun, wind, heat, pression, etc.) and converting it to electrical energy power [3, 4]. Such kind of networks is called energy harvesting wireless sensor network (EH-WSN)[5, 6].

Dahia et al. [7] developed a Markov-chain model for a sensor network in order to investigate the system performance in terms of energy consumption, data delivery and network capacity. There are also works that take into consideration retrial phenomenon involving the unreliability of the sensors, as it can be seen in the work of S. Zhang-Song et al [8]. They propose a model based on GSPN formalism to predict the energy consumption and introduce a sleeping mechanism to construct the energy plan. Wuechner et al. [9] propose the concept of unreliable orbit over which they construct a GSPN model to evaluate the performance of the WSN. Gharbi and Charabi [10] adopt an algorithmic approach based on GSPNs for modeling and analyzing finite-source WSNs with retrial phenomenon [11] and two servers classes. Boutoumi and Gharbi [12] propose an approach they call the two thresholds working vacation policy. They use a GSPN model that describes an energy saving and latency efficiency technique for full-duplex wireless sensor nodes.

Our paper aims to contribute to these ef-

forts that try to model and simulate WSNs by considering actual circumstances. More precisely, we propose a GSPN based formulation that models the energy consumption of the sensor due to message sending and investigate the influence of retrial messages, energy harvesting and sensor's state on the mean response time and energy efficiency of the Energy Harvesting wireless sensor network (EH-WSNs).

The rest of this paper is organised as follows: In section 2, we give a short description of the GSPN formalism. In section 3, we present a description of EH-WSNs. In the next section, we develop a GSPN model for EH-WSN and we define some performance formulas. In section 5, we derive and discuss numerical results by using TimeNet tool. Finally, we draw a conclusion as well as some directions for future works.

2 GSPN formulation

Generalised Stochastic Petri Nets are a powerfull graphical tool for modelling dynamic systems. GSPN are used in many areas such as concurrent programming and client-server architectures, to cite a few. These areas are characterised as containing parallel, asynchronous, distributed and stochastics problems[13]. A GSPN can be described by a bipartite graphs that consists of places and transitions. Places are represented by circles, whereas transitions are represented by bars. There are two kinds of transitions: immediate transition (represented by a thin black bar) which do not need a time to fire and timed transition (usually represented by boxes) which involves a delay until the firing becomes possible. Places in a GSPN can contain marks (represented by big dots). A mark (also called token) represents a condition of transition firing. Also, an inhibitor arc (denoted by a circle-head) is an arc which interdicts the firing of the associated transition. We refer interested readers to [14, 15] for more details.

Hence, to evaluate the performance of a Petri net, we need to derive the reachability graph . Then, we conduct a study of ergodicity [16] to verify the existence of a unique stationary state upon which we can compute several performance parameters such as: the mean number of tokens per place, the mean sojourn time in a place and the resource utilisation ratio, to name but a few [17].

3 GSPN model for EH-WSN

In this section, we propose a Petri net model to evaluate and analyse the performance of an EH-WSN by using a sensor-to-neighbor relationship abstraction. A sensor senses incidents from the vicinity and tries to send a report to the base station by sending a message to an idle neighbour. In its turn, a sensor node can receive messages from its neighbours and takes care of forwarding them further towards the base station. Figure 1 illustrates our GSPN model that considers retrial messages, energy consumption, energy harvesting, breakdowns and repairs.

At first, incidents are represented by tokens in the place $Msgs$. When a message arrives to *Attempt* after the transition *Arrival* firing, the message will enter the sending process if the sensor is active and has at least one free neighbour. Otherwise, the message joins the *Orbit* to call back later. After entering the sending process, a sufficient amount of energy (one token of the place *Battery*) will be consumed to complete the transmission of the message ($Send\_msg$ transition firing). The transition *working* firing represents the energy consumption for a normal operation. If the energy reaches the threshold $l1$, the sensor goes to *Standby* state immediately. The energy recovery process is represented by the periodic firing of the transition $Harvesting$.When the energy (number of tokens in the place *Battery*) reaches the threshold $l2$, the sensor leaves the standby state for the active state (firing of the transition Be_awake). However, a neighbour may fail during the sending process (firing of the transition $Breakdown$). The message returns to the orbit and a repair (or wait for an available neighbor) process is triggered (transition *Repar*) in order to revive a neighbor. Table 1 contains the characteristics of the timed transitions.

We can prove with a structural analysis that the GSPN model of Figure 1 is *bounded* and that the associated reachability graph is

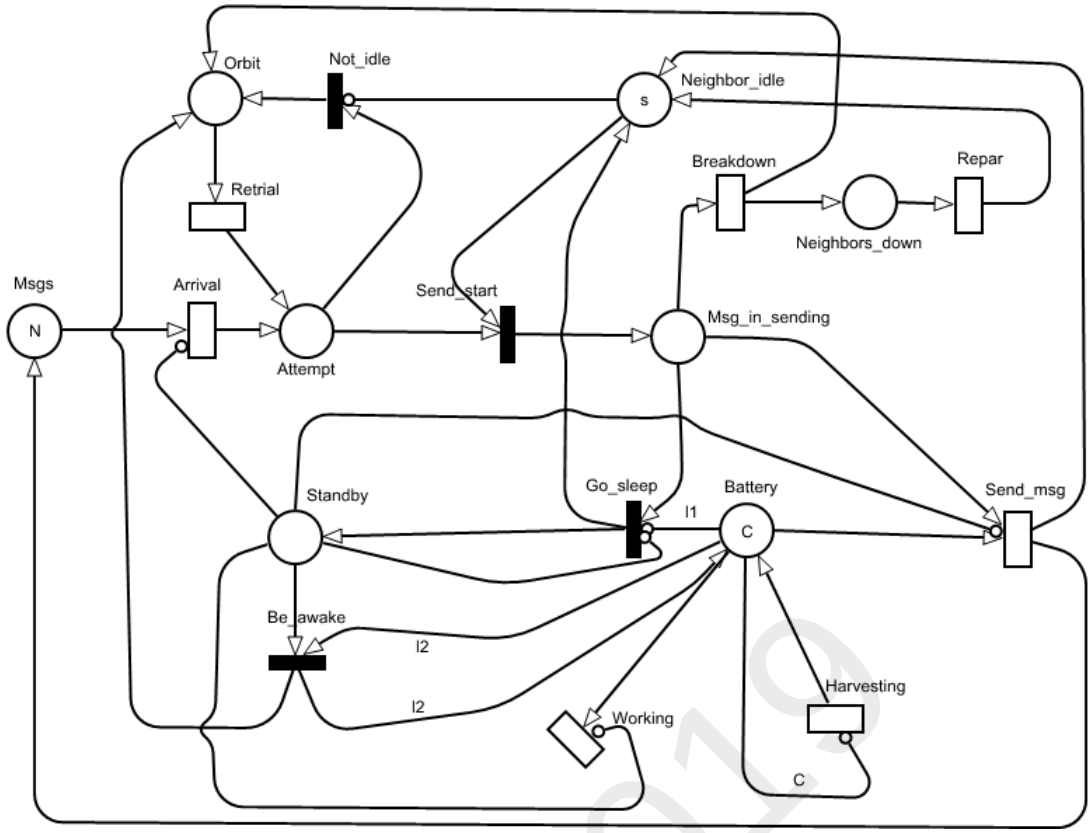


Figure 1: GSPN model for traffic communication in EH-WSN

strongly connected. Therefore, the model is ergodic and involves the existence of a unique steady state distribution. By solving the system assiciated to the steady state, several formulas of performance measures and reliability indexes can be derived. Applying Little's Law (The mean number of customers in a place is equal to the mean interarrival rate into it multiplied by the mean delay to traverse it) and other rules cited and proved in [17, 19], we can donate the following examples:

- The mean number of energy levels given by the mean number of tokens in the place $Battery$:

$$\overline{Battery} = \sum_{i:M_i \in M} M_i(Battery).\pi_i \quad (1)$$

- The mean number of retrial messages obtained from the mean number of marks in the place $Orbit$:

$$\overline{n_o} = \sum_{i:M_i \in M} M_i(Orbit).\pi_i \quad (2)$$

- The mean number of messages in the sending state given by the mean number of marks in the place $Msg\_in\_sending$:

$$\overline{n_s} = \sum_{i:M_i \in M} M_i(Msg\_in\_sending).\pi_i \quad (3)$$

- The throughput of message arrivals deduced from the debit of the transition $Arrival$:

$$\overline{\lambda} = \sum_{i:M_i \in M(arrival)} \lambda.M_i(Msgs).\pi_i \quad (4)$$

- The mean response time of a message that corresponds to the time between the arrival and the sending end.

$$\overline{R} = \frac{\overline{n_o} + \overline{n_s}}{\overline{\lambda}} \quad (5)$$

where M and π are the markup function and probability vector of steady state respectively.

Timed transition	Signification	Firing rate
Arrival	Arrival of a message	λ
Send_msg	Successfull sending of a message	μ
Retrial	Retrial of a message	ν
Harvesting	Energy recovery	ω
Working	Working energy consumption	γ
Breakdown	Failure of an active neighbour	δ
Repair	Repair of a neighbour	α

Table 1: Timed transitions description

Inputs	N	s	C	$l1$	$l2$	λ	μ	ν	ω	γ	δ	α
values	20	3	20	5	10	15	20	10	15	5	10^{-4}	1

Table 2: Inputs values

4 Numerical results

By the use of TimeNet tool[18] and the values of Table 2, we can find the follow graphs.

Figure 2 shows the influence of the harvesting rate on the battery mean charge level. When we vary the harvesting rate, the battery charge is between 4 and 16 unit levels . We can notice an exponential growth of battery charge if the harvesting rate is between 15 and 50. We conclude that the mean of battery charge enhances when the harvesting rate grows. This result proves that our model is compliant with an actual battery behavior.

Figure 3 shows the influence of message sending rate on the mean battery charge. When this rate increases, the mean battery charge decreases. It is clear that the percentage of battery charge drops dangerously when the sending rate exceeds 15 messages per time unit. That is, when the number of sent messages increases, the energy consumption increases. As a result, message transfer consumes the highest amount of energy in comparison to other activities in the network.

Figure 4 illustrates the impact of retrial attempts on the mean response time which is an important performance parameter. When the attempt rate increases i.e. the time between two attempts decreases, the mean response time decreases. This behaviour stems from the fact that message sending resumes shortly after the release of a neighbor or the activation of the main sensor.

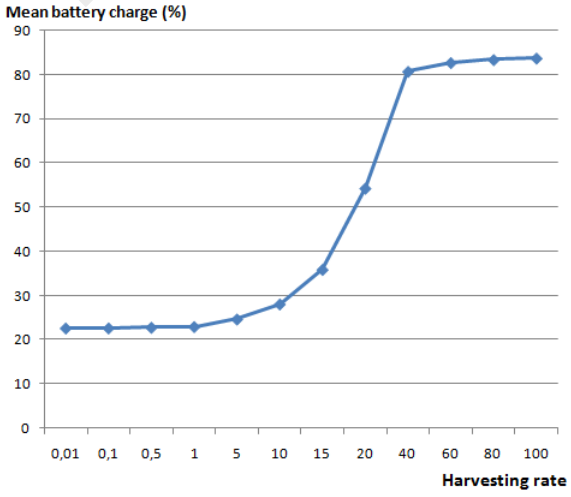


Figure 2: Mean battery charge versus harvesting rate

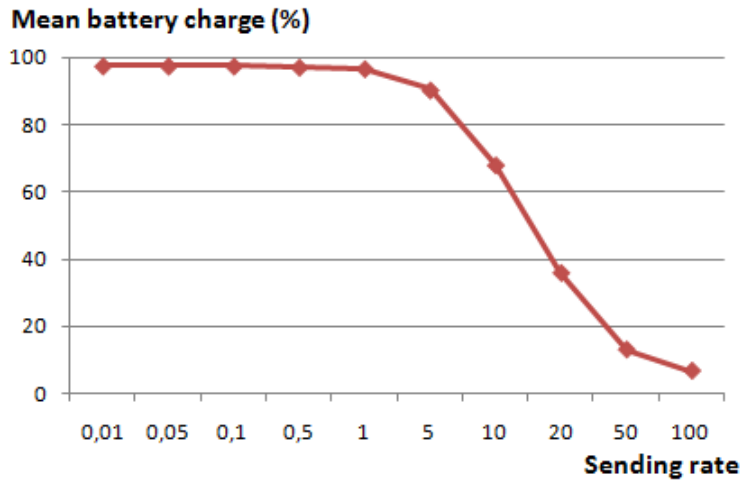


Figure 3: Mean battery charge versus sending rate

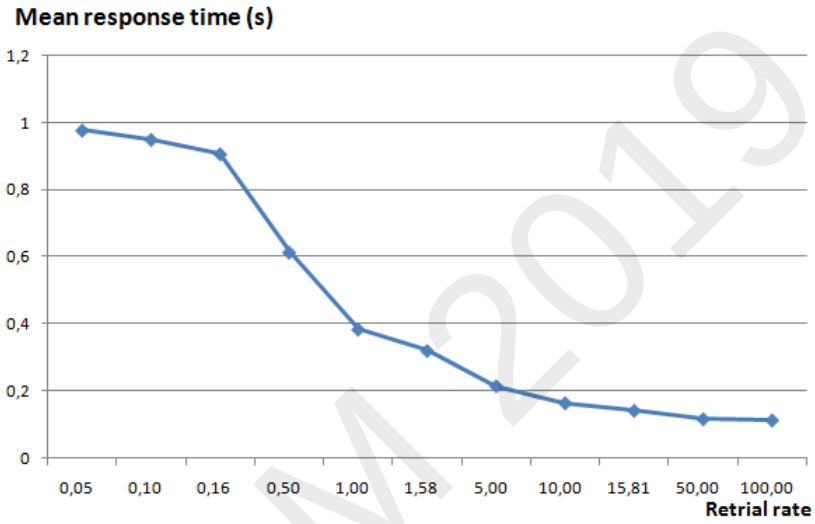


Figure 4: Mean response time versus retrial rate

5 Conclusion

In this paper, we propose an energy-aware model for wireless sensor networks. Our model is a GSPN formulation that takes into account message retrial, breakdowns, repairs and battery charging. The proposed model proves to be able to decide which input parameter to adjust to get a good trade-off between performance and contnuity of service.

In our future works, we focus on extendining our model in order to get more closer to reality by removing simplifying assumptions such as taking into account the differences between message types and the required energy to send them to the neighbors.

References

[1] Ian F Akyildiz, Weilian Su, Yogesh Sankarasubramaniam, and Edral Cayirci, Wireless sensor networks: a survey, Computer networks, 2002.

[2] Jacques Bahi, Wiem Elghazel, Christophe Guyeux, Mourad Hakem, Kamal Medjaher, and Noureddine Zerhouni. Reliable diagnostics using wireless sensor networks. Computers in Industry, 104:103 115, 2019.

[3] Wenqi GUO and William M. HEALY, Power Supply Issues in Battery Reliant Wireless Sensor Networks: A Review , International journal of intelligent control and systems vol. 19, NO. 1, MARCH 2014.

[4] Stefano Basagni, M Yousof Naderi, Chiara Petrioli, and Dora Spenza. Wireless sensor networks

with energy harvesting. Mobile Ad Hoc Networking: Cutting Edge Directions, pages 701–736, 2013.

[5] Faisal Karim Shaikh and Sherali Zeadally. Energy harvesting in wireless sensor networks: A comprehensive review. Renewable and Sustainable Energy Reviews, 55:1041 – 1054, 2016.

[6] Nouman Ashraf, Muhammad Faizan, Waqar Asif, Hassaan Khaliq Qureshi, Adnan Iqbal, and Marios Lestas. Energy management in harvesting enabled sensing nodes: Prediction and control. Journal of Network and Computer Applications, 132:104 117, 2019.

[7] R.Dahiya, A.K.Arora and V.R.Singh, Modelling the Energy Effcient Sensor Nodes for Wireless Sensor networks, J. Inst. Eng. India Ser. B(July-September 2015) 96(3):305-309 DOI 10.1007/s40031-014-0149-1.

[8] Zhang-song Shi, Cheng-fei Wang, Peng Zheng, and Hang-yu Wang. An energy consumption prediction model based on gspn for wireless sensor networks. In 2010 International Conference on Computational and Information Sciences, pages 1001–1004. IEEE, 2010.

[9] P. Wuechner, J. Sztrik, H. De Meer, Modeling wireless sensor networks using finite-source retrial queues with unreliable orbit, in: Proc. of the Workshop on Perf. Eval. of Computer and Communication Systems (PERFORM'2010), vol. 6821 of LNCS Publisher: Springer-Verlag, 2011.

[10] GHARBI Nawel et CHARABI Leila. Wireless networks with retrials and heterogeneous servers: Comparing random server and fastest free server disciplines. International Journal on Advances in Networks and Services Volume 5, Number 1 and 2, 2012, 2012.

[11] J.R. Artalejo, A classified bibliography of research in retrial queues: Progress in 1990–1999, Top 7 (1999) 187–211.

[12] Bachira Boutoumi and Nawel Gharbi, Two Thresholds Working Vacation Policy for Improving Energy Consumption and Latency in WSNs , Queueing Theory and Network Applications book, June 2018.

[13] M. Ajmone Marsan, G. Balbo, G. Conte, S. Donatelli, G. Franceschinis, Modelling with Generalized Stochastic Petri Nets, John Wiley and Sons, New York, 1995.

[14] Murata, T, Petri Nets : Properties, Analysis and applications, Proceedings of the IEEE, 77(4) (1989) 541–580.

[15] Peterson, J. L, Petri Net Theory and the Modeling of Systems, Prentice-Hall, 1981.

[16] Bacelli,F., Ergodic theory of stochastic Petri networks, INRIA Research Report No 1037, May 1989.

[17] G. Florin, C. Fraize, S. Natkin, Stochastic Petri nets: Properties, applications and tools, Microelectronics Reliability, Volume 31, Issue 4, Pages 669-697,1991, Pages 669-697.

[18] Armin Zimmermann , Modelling and Performance Evaluation With TimeNET 4.4, Quantitative Evaluation of Systems - 14th Int. Conf. (QEST 2017), Berlin Germany, September 5-7, 2017.

[19] Armin Zimmermann ,Stochastic Discrete Event Systems Modeling, Evaluation, Applications,Book: ISBN 978-3-540-74172-5, Springer-Verlag Berlin Heidelberg 2008.

Classification-based instance selection for Case-Based Reasoning

Abdelhak Mansoul
Dept. of computer science
University 20 Août 1955,
Skikda, Algeria.
LIO Laboratory, University
of Oran 1 Ahmed Ben Bella,
Oran, Algeria
a.mansoul@univ-skikda.dz

Baghdad Atmani
LIO Laboratory,
University of Oran 1 Ahmed
Ben Bella,
Oran, Algeria
atmani.baghdad@gmail.com

***Abstract*—**Case-Based Reasoning (CBR) is a problem solving technique that is attracting increasing attention. It solves new problems by matching and adapting the important features on the old cases that were successfully solved before. An important issue is how to select the best old case if more than one match are available, and how to perform efficiently the matching process by applying different techniques to avoid the adaptation step. In this work, we present a solution that integrates CBR and classification. The classification gives a reduced search space composed of representative cases called "Prototypes", instead of classes. Then, CBR uses the reduced search space for searching a current problem solution. Thereafter, we experiment our approach by using real-life datasets from UCI Machine Learning Repository, i.e., the iris plant dataset is employed for the benchmark test.

Keywords—Data mining, Classification Case-Based Reasoning, CBR, Decision

I. Introduction

Case-based reasoning (CBR) is an area of machine learning research that usually remind similar situations and recall their solutions. In other words, previous cases are applied to current problems by an analogical reasoning process for searching a solution. It has been extensively applied in different areas to support decision and was tested appropriately in different situations. It was widely applied to solve problems and support decision in health care or other domains [1], [2], [3]. However, it presents some shortcomings in its preliminary steps: the retrieval and the adaptation [1].

As a major shortcoming that will be cited is when the process found several similar cases and consequently several solutions a selection of an appropriate solution must be done. This operation is generally difficult and needs special procedures. So, different strategies have been proposed for the retrieval task. These strategies were using separately non-sequential indexing, nearest neighbor matching [4], ontologies [5], etc., or combined with other techniques : it's called the *Multimodal Reasoning*. It also combines data mining methods [6] as classification [7], [8], and clustering [9]. Multi-Criteria Analysis (MCA) [10] was also used and Rule Based Reasoning (RBR),..etc. The interest in multimodal approaches has reached different areas [6], [11], [12], [13]. It became an issue of current concern in CBR research [2], [3].

In this work, we are interested in investigating a new approach by integrating CBR and the classification method.

So, by a classification method we divide the whole case base in many sub groups or classes, and necessarily the case that may be close to the new case (new problem) will be in one and only one group. Thus, we will limit the search space to relevant cases.

The rest of the paper is structured as follows: Section 2, presents the main notions related to our topic. Section 3, gives a survey on some related works and underline the use of CBR with data mining techniques. Section 4, illustrates our approach using Classification-Based Instance Selection. Section 5, discusses the experimental results. Finally, Section 6 ends with concluding remarks.

II. Literature Review

CBR. It uses the principle of reusing past experiences. It is a powerful and frequently applied approach to solve problems. It is based on four tasks: retrieve, reuse, revise and retain. According to Aamodt and Plaza [14] there are four steps that a case-based reasoner must apply:

- retrieve similar cases;
- reuse a solution suggested;
- revise or adapt the solution for the new case;
- retain the new validated solution.

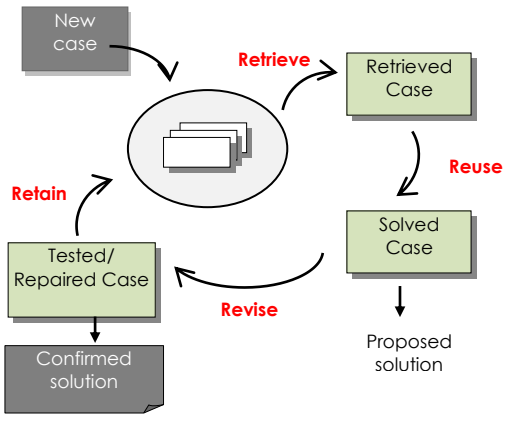


Fig. 1. The CBR cycle [14]

Classification. It is a data mining task. It is the problem of identifying to which of a set of categories (sub populations), a new item belongs to. Classification consists of assigning a class label to a set of unclassified items. The classification uses some techniques such as [15]:

- decision Tree based Methods;
- rule-based Methods;
- nearest-neighbors;
- neural networks;
- naïve bayes and bayesian networks;
- support vector machines.

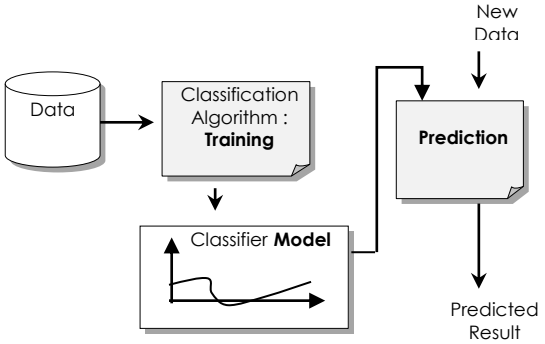


Fig. 2. The classification process

Benefits of classification techniques to CBR.

As CBR looks for the solution on a large number of cases, it is more interesting to reduce this search space. Indeed, and as classification operates by the division of a space containing objects in several spaces (groups) rather reduced, it will be more interesting to be able to use this technique and then to locate the best subspace as small as possible which could contain the desired solution.

III. Related Work

There are several studies on use of CBR. Only, there is discrepancy between theory and experimental approaches, There are also many successful works concerning the CBR approach in different domains [2], [3], [16].

Combining CBR with other reasoning methods (multimodal reasoning) as MCA [3], [10], [17], Data mining [4], [9], etc., has also caught a lot of attention. Multimodal reasoning aims to provide a better way for searching solutions. Previous research in this area has been undertaken by various researchers [2], [3], [5], [7].

Data mining techniques were used in different manners to facilitate CBR mainly by using association rules, classification and clustering [4], [6].

1) Combining with Association Rules

Reasoning with rules was the first approach to be successfully integrated with CBR, since rules are exploited to make decisions. Thus, rules were well integrated with CBR and many solutions were proposed.

Verma et al., proposed a solution based on CBR-RBR which will eventually help provide the convenient solution to the given problem depending upon the business [18]. Cabrera and Edye used CBR and RBR to diagnose acute bacterial meningitis [12]. Saraiva et al., applied RBR to improve the CBR's retrieval process [19].

Bilgi et al., developed a symptomatic decision support system for Neurological Disorders. This system uses rules of the neurology domain and a framework to learn from the cases of the patients. The system will give a symptomatic diagnostic conclusion [20].

Balakrishnan et al., proposed a retinopathy prediction system based on association rules and CBR [11].

Ting et al., propose a new revised CBR, named Rule-Associated CasE-based Reasoning (RACER), which integrates CBR and association rules mining for supporting General Practitioners prescription [21].

Pandey and Kundra developed a medical decision support for the diagnosis of electroencephalography (EEG)-based diseases, integrating J48 (decision tree) and CBR. The data mining method is used for reducing the dimension of parameters and CBR is used for diagnosis of the different EEG-based diseases [9].

2) Combining with Classification

Vikas et al., developed a system based on CBR and classification of hepatitis C virus (HCV) diagnosis. The classification is used to predict the presence of hepatitis C virus which helps in early diagnosis of the infection [8].

Ramos-González and al., presented a CBR-Classification framework applied to squamous cell carcinoma and adenocarcinoma discrimination, aiming to provide accurate diagnosis [22].

3) Combining with Clustering

Schmidt et al. suggested clustering cases into prototypes and remove redundant ones to avoid an infinite growth of case base, the retrieval searches only among these prototypes [23].

Vong et al. used the combination CBR-Clustering to illustrate the efficiency of CBR and clustering in automotive engineering diagnostic problem [24].

IV. The Classification-Based Instance Selection for CBR

The case we consider is defined by a set n of features F and a solution S considered in the case. Thus, this case will be described as follows:

[Case_Instace]
[Feature$_1$] …., [Feature$_n$] [S:Solution]
[End_Case_Instance]

A. The contribution of Classification-Based Instance selection.

The framework is based on the notion of Instance Selection. This notion assumes that a single case is able to

represent information of an entire subset of cases. In general, this notion is known as : Prototype. Prototypes leads to find solutions on a small set of items instead of the whole case base.

Instance Selection results in a subset of the original set of items, and is often based on notions of "prototypicality". For each item in the original subset of items, "prototypicality" measures the degree of representativeness of this item among all Items in the subset [25], [26] .

In our case we use the simple notion of barycenter as "prototypicality" measure or representativeness of a subset. Thus, the element in the barycenter is the representative of the sub-class. By this idea, we intend to find the more representative case of the subset. Thus, this strategy reduces the search space. However, we can use more than one prototype for sub-class to have more items in the reduced case base. Thus, for each class deduced from the whole case base, we can have *m* prototypes.

B. The framework (CBI4CBR)

All the processing will be handled by the following units:

1) Classification-Based-Instance Selection (CBI)

The key idea to use classification technique, for instance selection purposes uses the notion of similarity and consists of involving all features of a case instead of sorting case simply on the last feature (class) which is not a robust classification but a simple sorting of cases.

The set of representative cases will form the output of Cassification-Instance-Selection.

- *Classification.* Initially a case-based classification method is used to pre-process the cases. For the purposes of our experiment, we use *k-means* algorithm [15], for its easy processing. An application of a classification draws the initial classes ($Class_1$, $Class_2$, …, $Class_n$).
- *Search for Relevant Instances.* At this step, searching relevant instances **(**prototypes) is done in each class found ($Class_i$), to make a reduced space search for the next step (*CBR*). This allows the selection of a smaller set of representative cases which constitutes the final output of Search for Relevant Instances.

2) Case-Based Reasoning (CBR)

CBR launch respectively its different steps (retrieve, reuse, revise, retain) using as input the Reduced Case Base, founded at unit 1 (*Classification-Based-Instance*) for deducing the relevant solution.

This whole process will be performed the following framework.

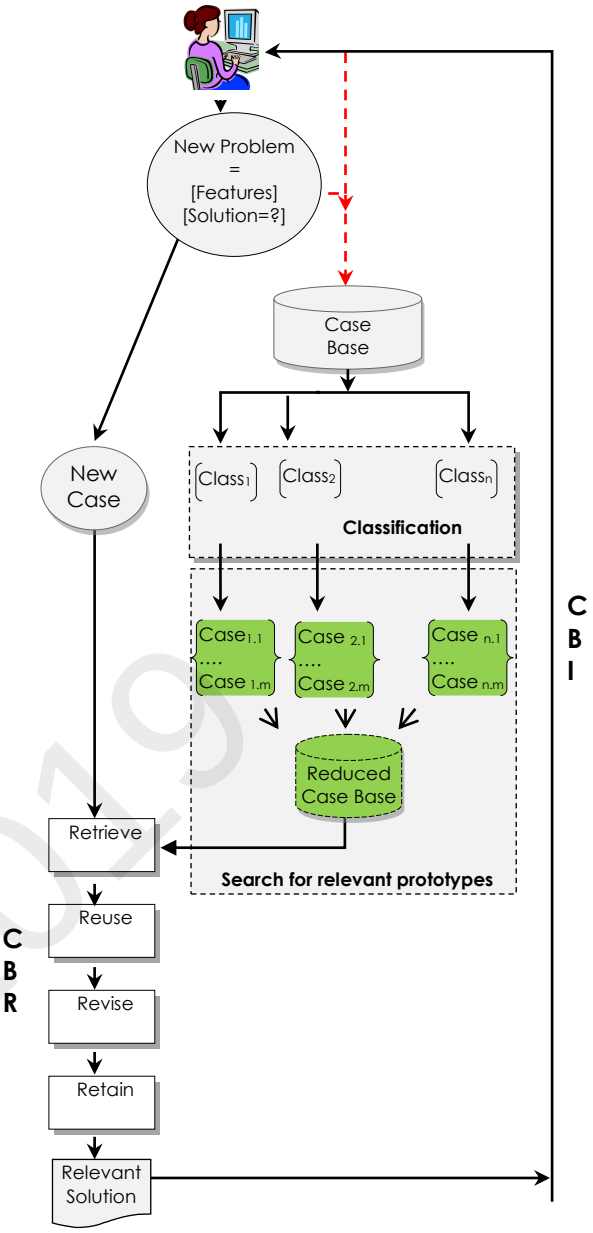


Fig. 3. Overview of the framework *Classification-Based-Instance-for-CBR (CBI4CBR)*

C. The CBI4CBR processing

- *CBI*

Reducing the space of similar cases is an essential operation before CBR's retrieval step. The reducing space of similar cases can clearly make retrieval of the best solution computationally better regarding to only relevant cases (interesting solutions) for the current case being processed and more meaningful since only relevant cases are considered. All this processing is done sequentially by the three steps below:

(1) Initiate classification: For each iteration, save the new barycenter and its class.
(2) Search for relevant cases: For each class, select *m* (*m*>=1) relevant cases.
(3) Save relevant cases in the Reduced Case Base.

These steps will be performed by the following pseudo-algorithm.

Pseudo-algorithm: *Classification-Based-Instance* ***(CBI)***

```
Input:    Case-Base, m : number of instances selected per class,
Output:   Reduced-Case-Base
1: Begin
2: Initialize k for k-means
3: Initilize m to select relevant cases
4: Perform k-means (Case-Base):
5: For each Iteration of k-means do
  {
     For Current_BaryCenter
      { Save Current_BaryCenter in
                          Vector-BaryCenter-Of-Classi;
        Save Class-Of-Current_BaryCenter in
                           Vector- Class-Of-Current_BaryCenter;
        Save Order_Of_Current_BaryCenter in
                          Vector-Order_Of_Current-BaryCenter
      }
  }
  EndFor
6: For Each Vector-BaryCenter-of-Classi ( i=1: m) do
  Prototypes_Founded = Search Prototypes in
                          Vector-BaryCenter-Of-Classi
    Prototypes-Classi = Prototypes-Founded
    Insert Prototypes-Classi In Reduced-Case-Base
    Next i
  EndfFor
7: End
```

- *CBR*

It consists in finding the *k* closest cases of the proposed case by using a similarity measure. The *k-nn* method is used for its easy implementation. The process will select similar cases and will extract the preliminary *Relevant_Solution* that will be considered for the current case. This process will be handled by the following pseudo-algorithm.

Pseudo-algorithm : ***CBR***

```
Input:      Reduced-Case-Base, New-Case=(F,S=" ")
Output:     Relevent-Solution
1: Begin
2: Classification_Based_Instance()
3: Cases-To-Reuse = Retrieve(New-Case, Reduced-Case-Base)
4: Reuse(Cases-To-Reuse, Adapted-Solution)
5: Revise(Adapted-Solution, Confirmed-Solution)
6: If Confirmed-Solution is accepted Then
     Relevant-Solution := Confirmed-Solution
     Retain_new_case(F, Relevant-Solution)
   Endif
7: End
```

V. Experimental Setup

The proposed approach has been applied to Iris Plants Database [27]. We project to use the four features to classify iris plants. The framework was developed as an application in Java using libraries of JColibri [28] to perform CBR's operations and Weka APIs for classification [28].

a. *Data Description*

We use the Iris Plants Database. Table 1 gives an overview of data sample uploaded from UCI Machine Learning Repository [27]:

TABLE 1. IRIS PLANTS DATABASE SAMPLE [27].

Data Sample
5.1,3.5,1.4,0.2,Iris-setosa
6.8,2.8,4.8,1.4,Iris-versicolor
6.7,3.1,5.6,2.4,Iris-virginica

"Each iris plant is represented in the data set by four features and a predicted feature : class of iris plant." [27].

The following labels are used in the class: "Iris-setosa", "Iris-versicolor", "Iris-virginica".

- sepal length in cm
- sepal width in cm
- petal length in cm
- petal width in cm
- class: (Iris Setosa, Iris Versicolour, Iris Virginica)

For the purposes of our experimentation we have transformed the Iris Plants Database into a case base and each case will be described by the descriptive feature F_1, F_2, F_3, F_4, and have noted *S* for the class. Table 2 shows case base features.

TABLE 2. CASE BASE FEATURES

Features	Label
F_1	sepal length in cm
F_2	sepal width in cm
F_3	petal length in cm
F_4	petal width in cm
S	Class (Iris Setosa, Iris Versicolour, Iris Virginica)

b. *Creation of Iris Plants case base, leaning and testing case bases : Iris, $Iris_L$, $Iris_T$*

Initially, we transform the Iris Plants Database into Iris Plants Case Base *Iris*. It contains 150 Instances of plants. IrisCaseBase will be described by the features F_1, F_2, F_3, F_4, previously defined. For each case ω_i we associate a target attribute denoted S, which takes its values in the set {"Iris-setosa", "Iris-versicolor", "Iris-virginica"}. Table 3 shows cases noted $\omega_1, \omega_2, \ldots, \omega_n$ of Iris Case Base.

TABLE 3. IRIS

ω	$F_1(\omega)$	$F_2(\omega)$	$F_3(\omega)$	$F_4(\omega)$	$S(\omega)$
ω_1	5.1	3.5	1.4	0.2	Iris-setosa
ω_2	6.8	2.8	4.8	1.4	Iris-versicolor
…					
ω_n	6.7	3.1	5.6	2.4	Iris-virginica

Then we split Iris Case Base as follow :

$IRIS_L$=60 % of IRIS and $IRIS_T$=40% of IRIS.

c. *Experimentation*

We consider an individual draw from the testing base $Iris_T$. At each draw we assume that the case has an original class that we don't know, and it's the final comparison with the result of the system through $Iris_L$, on which will establish us that the solution found is equivalent or not to the original class.

We repeat the draw as many times until we have selected 20 successive cases for the same class from $Iris_T$. At each iteration of the draw the following heuristic is checked:

This treatment will be repeated as many existing classes.

$\forall\ \omega_i \in Iris_T$ and $\forall\ \omega_j \in Iris_L$

{ If [$F(S(\omega_i) = F(S(\omega_j))$] Then *well-match*
Else *mismatch* }

d. *Evaluation*

We calculate the error rate of each class using the formula (1). This rate represents the number of cases differently mismatched (in term of classes) in $Iris_L$ compared to the original class in $Iris_T$. The test results are presented in Table 4.

$$Error\ Rate = (MC * 100) / C \qquad (1)$$

MC : Total of mismatched cases

C : Total cases tested from $Iris_T$

e. *Results*

As presented in Table 4, we can notice that the rate of well matching (similar class) is relatively important.

The accuracy of *CBI4CBR* model is 78% for all tests, that means it is relatively high and provides hopeful results for discovering a class close the reality as existing in testing case base $Iris_T$. This result could help CBR to compute effective solutions.

TABLE 4. RESULTS WITH ERROR RATE (ER)

Cases tested	Tested cases Class	well-match	mismatch	Error Rate %
20	Iris Setosa	17	3	15
20	Iris Versicolour	14	6	30
20	Iris Virginica	16	4	20

VI. CONCLUSION AND FUTURE TRENDS

Multimodal reasoning has shown more interesting result with CBR through the different study done by researchers. Only, the different results obtained need improvements. In this context, our study tries to address a weakness of CBR.

Future orientation of our work intends to evolve our approach towards the following points :

- a study, which allows the decision maker to define efficiently the number of prototypes needed for testing ;
- a study for choosing the most interesting features that allow searching efficiently classes solutions.

This orientation can compute efficiently a Search Space Reduction and also the final solution to a new problem.

REFERENCES

[1] S. Begum, M. U. Ahmed, P. Funk, N. Xiong, and M., Folke, "Case-Based Reasoning Systems in the Health Sciences: A Survey of Recent Trends and Developments," IEEE Transactions on systems, man, and cybernetics part c: applications and reviews, vol. 41 no 4, pp. 421-434, 2011.

[2] I. Bichindaritz, and C. Marling, "Case-based reasoning in the health sciences: Foundations and research directions," Computational Intelligence in Healthcare 4. Springer Berlin Heidelberg, pp. 127-157, 2010

[3] S. Montani, "Exploring new roles for case-based reasoning in heterogeneous AI systems for medical decision support," Applied Intelligence, vol. 28, no 3, 275-285, 2008.

[4] A. Nega, and A. Kumlachew, "Data Mining Based Hybrid Intelligent System for Medical Application," International Journal of Information Engineering and Electronic Business, 9(4), 38, 2017.

[5] F. Xu, X. Liu, W. Chen, C. Zhou, and B. Cao, "Ontology-Based Method for Fault Diagnosis of Loaders Sensors," 18(3), 729, 2018.

[6] I. Bichindaritz, "Data Mining Methods for Case-Based Reasoning in Health Sciences.," In ICCBR (Workshops), pp. 184-198. 2015.

[7] F. Mezera, and J. Krupka, "Classification of Clients on the Basis of Modifying Case-Based Reasoning Algorithms," In International Conference on Man–Machine Interactions (pp. 311-319). Springer, Cham, 2017.

[8] B. Vikas, D. V. S. Yaswanth, W. Vinay, B. S. Reddy, and A. V. H. VSaranyu, "Classification of Hepatitis C Virus Using Case-Based Reasoning (CBR) with Correlation Lift Metric," In Information Systems Design and Intelligent Applications (pp. 916-923). Springer, Singapore. 2018.

[9] B. Pandey, and D. Kundra, "Diagnosis of EEG-based diseases using data mining and case-based reasoning," International Journal of Intelligent Systems Design and Computing, 1(1-2), 43-55, 2017.

[10] N. Armaghan, and J. Renaud, "An application of multi-criteria decision aids models for Case-Based Reasoning," Information Sciences, Vol 210. pp. 55-66, 2012.

[11] V. Balakrishnan, M. R. Shakouri, and H. Hoodeh, "Integrating association rules and case-based reasoning to predict retinopathy," Maejo International Journal of Science and Technology, vol. 6, no 3. 2012.

[12] M. M. Cabrera, and E. O. Edye, "Integration of rule based expert systems and case based reasoning in an acute bacterial meningitis," clinical decision support system. arXiv preprint arXiv:1003.1493. 2010.

[13] B. Pandey, and D. Kundra, "Diagnosis of EEG-based diseases using data mining and case-based reasoning," International Journal of Intelligent Systems Design and Computing, 1(1-2), 43-55, 2017.

[14] A. Aamodt, and E. Plaza, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," AI communications, vol. 7 no 1, pp. 39-59, 1994.

[15] S. K. David, A. T. Saeb, M. Rafiullah, and K. Rubeaan, Classification Techniques and Data Mining Tools Used in Medical Bioinformatics. In *Big Data Governance and Perspectives in Knowledge Management* (pp. 105-126). IGI Global. 2019.

[16] S. A. Darabi, and B. Teimourpour, "A Case-Based-Reasoning System for Feature Selection and Diagnosing Asthma," In Handbook of Research on Data Science for Effective Healthcare Practice and Administration (pp. 444-459). IGI Global. 2017.

[17] A. Mansoul, and B. Atmani, "Combining Multi-Criteria Analysis with CBR for Medical Decision Support," Journal of Information Processing Systems, vol. 13, no 6, 2017.

[18] L. Verma, S. Srinivasan, and V. Sapra, "Integration of rule based and case-based reasoning system to support decision making," Issues and Challenges in Intelligent Computing Technics (ICICT). International Conference on. IEEE, pp. 106-108, 2014.

[19] R. Saraiva, M. Perkusich, L. Silva, C. Siebra, and A. Perkusich, "Early diagnosis of gastrointestinal cancer by using case-based and rule-based reasoning," Expert Systems with Applications, vol. 61, pp. 192-202, 2016.

[20] N. B. Bilgi, G. M. Wali, A. Mense, P. Takkekar, B. Yaddi, and S. Patil, "Symptomatic Decision Support System for Neurological Disorders," BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 8(4), 5-16, 2017.

[21] S. L. Ting, W. M. Wang, S. K. Kwok, A. H. Tsang, and W. B. Lee, "RACER: Rule-Associated CaseE-based Reasoning for supporting General Practitioners in prescription making," Expert systems with applications, 37(12), 8079-8089, 2010.

[22] J. Ramos-González, D. López-Sánchez, J. A. Castellanos-Garzón, J. F. De Paz, and J. M. Corchado, « A CBR framework with gradient boosting based feature selection for lung cancer subtype classification. Computers in biology and medicine, 86, 98-106, 2017.

[23] R. Schmidt, S. Montani, R. Bellazzi, L. Portinale, L. Gierl "Cased-based reasoning for medical knowledge-based systems". International Journal of Medical Informatics, Vol. 64 No 2. 355-367. 2001.

[24] C. M. Vong, P.K. Wong, and W.F. Ip "Case-based classification system with clustering for automotive engine spark ignition diagnosis." In: Computer and Information Science (ICIS), IEEE/ACIS 9th International Conference on. IEEE, p. 17-22.2010.

[25] D. B. Skalak, Prototype and feature selection by sampling and random mutation hill climbing algorithms. In *Machine Learning Proceedings*, pp. 293-301. Morgan Kaufmann. 1994.

[26] R. Schmidt,and L. Gierl, Case-based reasoning for antibiotics therapy advice: an investigation of retrieval algorithms and prototypes. *Artificial intelligence in Medicine*, *23*(2), 171-186. 2001.

[27] https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data

[28] ‚J. J. Bello-Tomás, P. A. González-Calero, and B. Díaz-Agudo, "Jcolibri: An object-oriented framework for building cbr systems," Advances in case-based reasoning. Springer Berlin Heidelberg, 32-46, 2004.

[29] G. Holmes, A. Donkin, and I. H. Witten, "Weka: A machine learning workbench. In Intelligent Information Systems," Proceedings of the 1994 Second Australian and New Zealand Conference on. IEEE, 357-361. 1994.

Numerical Analysis and Simulation of a two Dimensional Fractional Order map with Its Control

Amina Aicha Khennaoui
Departement of Mathematics and Computer Sciences
University of Larbi Ben M'hid
Oum El Bouaghi, Algeria
kamina_aicha@yahoo.fr

Adel Ouannas
Department of Mathematics and Computer Science
University of Larbi Tebessi
Tebessa, Algeria
ouannas.a@yahoo.fr

***Abstract*—In this paper a new fractionalized two-dimensional map is proposed based on the Caputo-like difference operator. Bifurcation and chaos for different values of the fractional order parameter are presented. Also, the largest Lyapunov exponent is calculated. Finally, chaos in the generalized map is controlled using the linearization method.**

Keywords—Discrete-time, fractional calculus, Hénon-Lozi map, numerical analysis

I. INTRODUCTION

Chaotic dynamical systems have been a subject of active research in several areas of science, such as physics, biology, ecology, etc. They can be divided into two major categories: continues-time and discrete-time systems. In recent years, the application of discrete fractional calculus has attracted extensive attention in the field of engineering and biological research due to the long memory effect. Wu and Balaneau [1] have introduced a new fractionalized map by applying discrete fractional calculus tools on an arbitrary time scale. The theory of delta difference equation was used to investigate general properties of chaotic behavior. In [2] a fractional generalization of the standard maps has been derived. A fractional generalization of the Hénon map was also suggested in [3]. The authors point out that the chaotic patterns exhibited by the fractional generalized Henon map depend on the fractional order. This means that the fractional map is more suitable for secure communications and encryption, as it includes a new degree of freedom.

In the current work, we examine the fractional map corresponding to the unified discrete-time system and study its dynamics and control.

II. FRACTIONAL FORMULA AND NUMERICAL SOLUTION

In this paper, we are interested in the two order difference equation [4], which is described as

$$\begin{cases} x(n+1) = 1 - 1.4 f_\alpha(x(n)) + y(n), \\ y(n+1) = 0.3x(n) \end{cases} \quad (1)$$

where x,y are the states variables, α is system parameter define in the range $\alpha \in [0,1]$, and the function f_a is given by $f_\alpha(x_n)=\alpha|x_n|+(1-\alpha)x^2_n$. There are three situations for the Unified map (1) which are based on the bifurcation parameter α. When α close to 0 the system behave like Hénon map second, and for α close to 1 the system behave like the Lozi map and when $0<\alpha<1$ the unified map (1) is chaotic with different types of attractors. The phase portraits of the chaotic maps for $\alpha=0$, $\alpha=1$ are shown in Fig.1-(a) and Fig.1(b) with an initial condition $(x_0,y_0)=(0,0)$. The largest Lyapunov exponents and bifurcation diagram are shown in Fig.2.

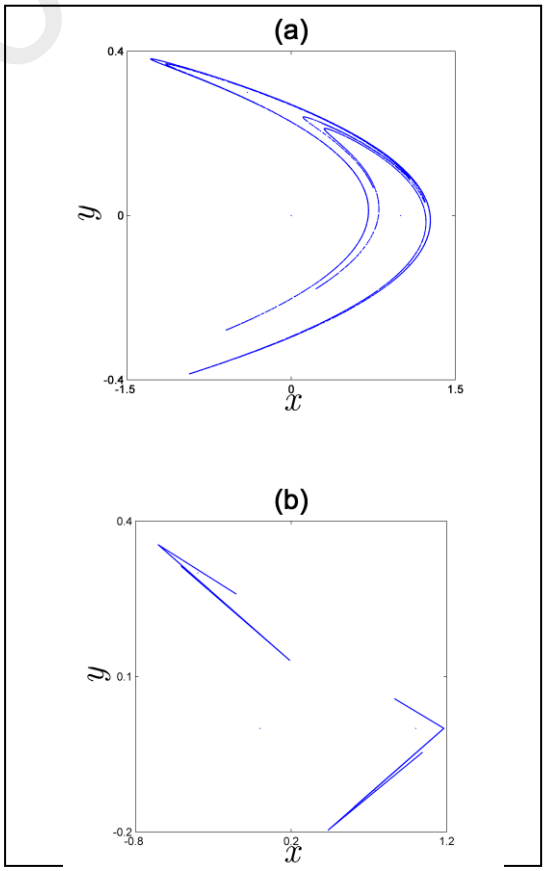


Figure 1: (a) chaotic attractor obtained for $\alpha = 0$, (b) chaotic attractor obtained for $\alpha = 1$.

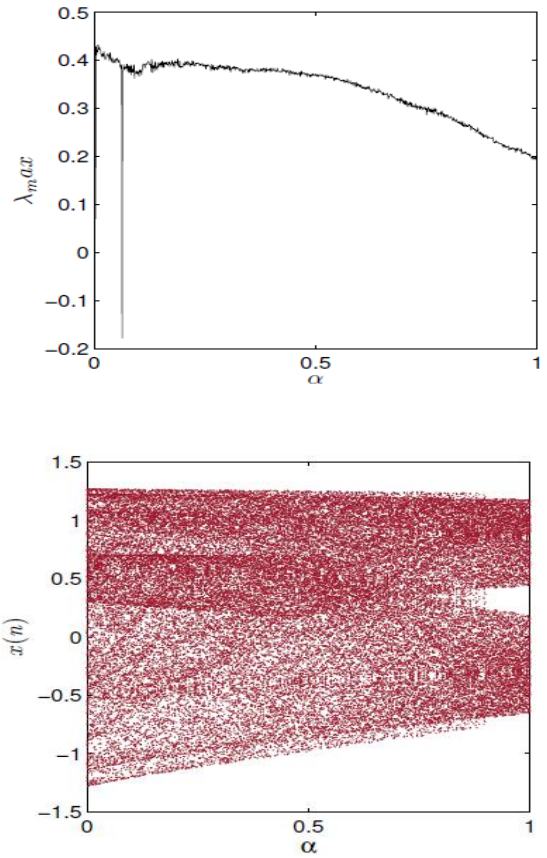


Figure 2: Bifurcation diagram and largest Lyapunov exponents for α as bifurcation parameter.

The main object of this paper is to define new fractionalized two dimensional map based on the Caputo-like difference operator. For that, we start by defining some basic concept of discrete fractional calculus. First, we define the *v*–th fractional sum for $v > 0$, as

$$\Delta_a^{-\nu} X(t) = \frac{1}{\Gamma(\nu)} \sum_{s=a}^{t-\nu} (t-s-1)^{(\nu-1)} X(s), \tag{2}$$

in particular, the *v*-fractional sum map every function defined on N_a to function defined on N_{a+v}. Here $t^{(v)}$ is the falling factorial function, defined as

$$t^{(\nu)} = \frac{\Gamma(t+1)}{\Gamma(t+1-\nu)}, \tag{3}$$

where Γ is the Euler gamma function.

In this paper we use one of the most common definitions in discrete fractional calculus the Caputo-like difference operator, to carry out the fractionalized formula. The Caputo-like difference operator for $n=\lceil v \rceil+1$, and $v>0$ is given as Eq.(4)

$$ {}^{C}\Delta_a^{\nu} X(t) = \Delta_a^{-(n-\nu)} \Delta^n X(t) = \frac{1}{\Gamma(n-\nu)} \sum_{s=a}^{t-(n-\nu)} (t-s-1)^{(n-\nu-1)} \Delta_s^n X(s), \tag{4}$$

where $t \in N_{a+n}$.

The fllowing theorem can be used to describe the numerical formula for fractional order difference equation.

A. Theoreme

For the fractional difference equation

$$\begin{cases} {}^{C}\Delta_a^{v} u(t) = f(t+v-1, u(t+v-1)), \\ \Delta^k u(a) = u_k, \ n = [v]+1, \ k = 0,1,...,n-1, \end{cases} \tag{5}$$

$$u(t) = u_0(t) + \frac{1}{\Gamma(v)} \sum_{s=a+n-v}^{t-v} (t-\sigma(s))^{(v-1)} f(s+v-1, u(s+v-1)), t \in \mathbb{N}_{a+n}, \tag{6}$$

where

$$u_0(t) = \sum_{k=0}^{m-1} \frac{(t-a)^k}{k} \Delta^k u(a). \tag{7}$$

Now, the fractional order map is obtained using the Caputo like difference Operator as followed:

$$\begin{cases} {}^{C}\Delta_a^{v} x(t) = 1 - 1.4 f_\alpha (x(t-1+v)) + y(t-1+v) - x(t-1+v), \\ {}^{C}\Delta_a^{v} y(t) = 0.3x(t-1+v) - y(t-1+v), \end{cases} \tag{8}$$

where a is the starting point and *v* is the fractional order, using *Theorem.A*, the numerical solution of system is given as follows:

$$\begin{cases} x(n) = x(0) + \frac{1}{\Gamma(v)} \sum_{j=1}^{n} \frac{\Gamma(n-j+v)}{\Gamma(n-j+1)} \\ \qquad (1 - 1.4 f_\alpha (x(j-1)) + y(j-1) - x(j-1)), \\ y(n) = y(0) + \frac{1}{\Gamma(v)} \sum_{j=1}^{n} \frac{\Gamma(n-j+v)}{\Gamma(n-j+1)} (0.3x(j-1) - y(j-1)). \end{cases} \tag{9}$$

From Eq.(9) we notice that the solution *x(n)* have memory effect means that it depends on all the previous states *x(n-1),x(n-1). ...x(1).*

III. DYNAMICAL ANALYSIS

In this section we use phase portraits bifurcation diagram largest Lyapunov exponents to show that the dynamical of the fractionalized map is affected by the changes of the fractional order v. Firstly, we fix the initial condition to (x0 ,y0)=(0,0), then we vary the fractional order v from 0 to 1. The bifurcation diagrams will provide general information about the dynamical behavior of the new map with the variation of parameter α,. Beside the Largest Lyapunov exponent will characterize the exponential rate of separation of infinitesimally close trajectories in phase space. A positive LLE is usually taken as an indication of chaos, provided the dynamical system is bounded. Fig.3 a,d Fig.4 shows the x-α bifurcation diagrams and largest Lyapunov exponent of the fractionalized map. From

the bifurcation diagrams, it can be observed that states of the system are different. When the values of ν are close to 1 the bifurcation diagram are similar to the integer order one see Fig.3. As ν decreases further, we notice that the interval where the chaos apparent get shrink, as can be seen in Fig.4 for ν=0.45, chaos is seen in the interval α∈[0.75,0.95]. Border-collision bifurcation scenario is observed for *ν=0.45*. The fractional map (8) begins with a fully-developed chaotic regime, and increasing *a* leads to the disappearance of the chaotic band and the appearance of a 8-period orbit. Phase portraits are shown in Figs. 5, from which we can see that the system exists periodic and chaos for different values of the parameter *a*.

IV. Control of the Fractional-Order Map

In this section, we are interested in a one-dimensional adaptive control law that forces the states of the proposed two-dimensional fractional unified system to zero asymptotically.

A. *Theorem:*

The 2D fractional-order unified chaotic map can be controlled under the 1D control law:

$$U = 1.4 f_{\alpha}(x(t)) - 1.3y(t) - 1. \tag{10}$$

Since the aim of the control law (10) is to force the two states towards zero asymptotically, what we want to do is to show that the zero solution of this resulting controlled system dynamics is globally asymptotically stable. In order to do so, we employ the fractional discrete Lyapunov method described in *Theorem A*. We propose the Lyapunov function:

$$ {}^{C}\Delta_a^{v} V(t) = \frac{1}{2}\left({}^{C}\Delta_a^{v} x^2(t) + {}^{C}\Delta_a^{v} y^2(t)\right). \tag{11}$$

In order to put Theorem.A to the test, a MATLAB script was run taking *v* = 0.95 and *a* = 0. The result in Fig.6 clearly shows how the states progress towards zero.

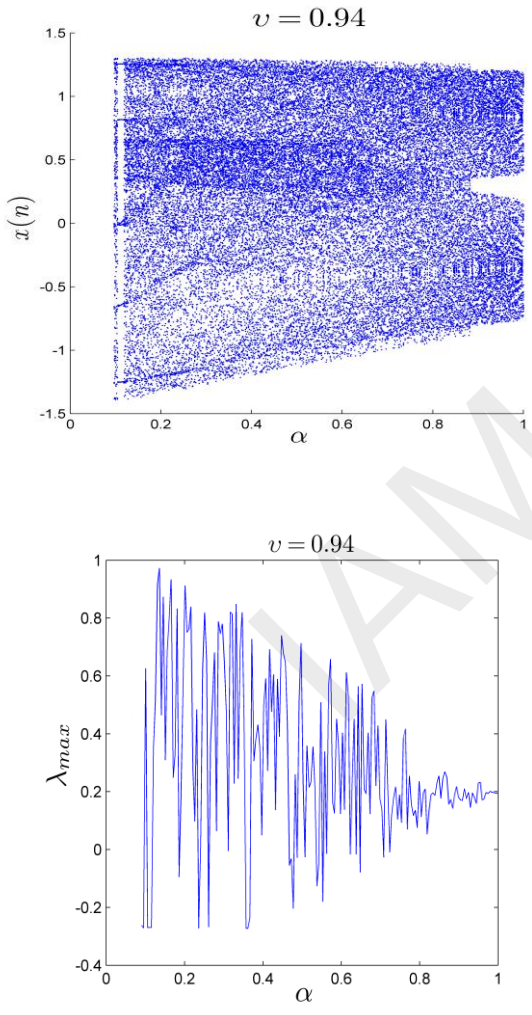


Figure 4:Bifurcation diagram and Largest lyapunov exponent of the fractionalized map (8) for *v=0.94*.

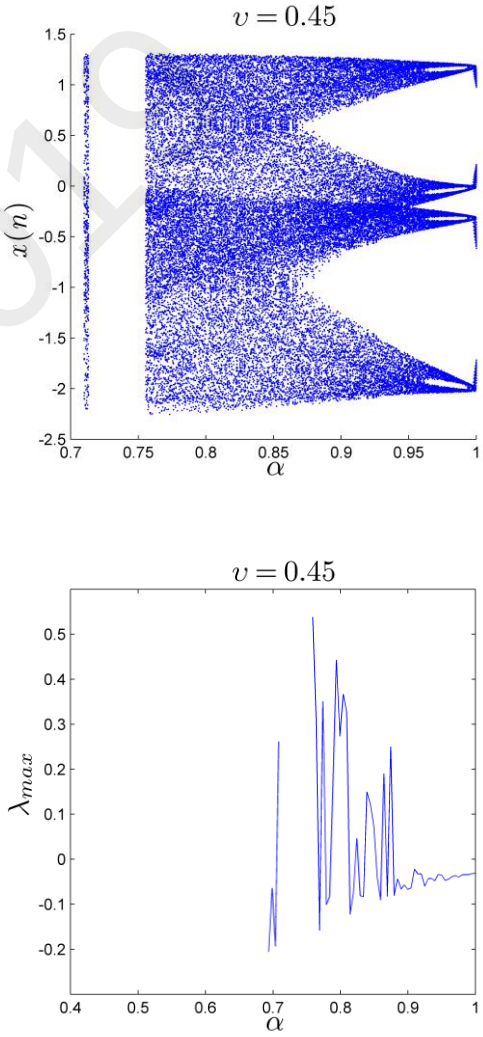


Figure 3: Bifurcation diagram and largest Lyapunov exponents of the fractionalized map for v=0.45

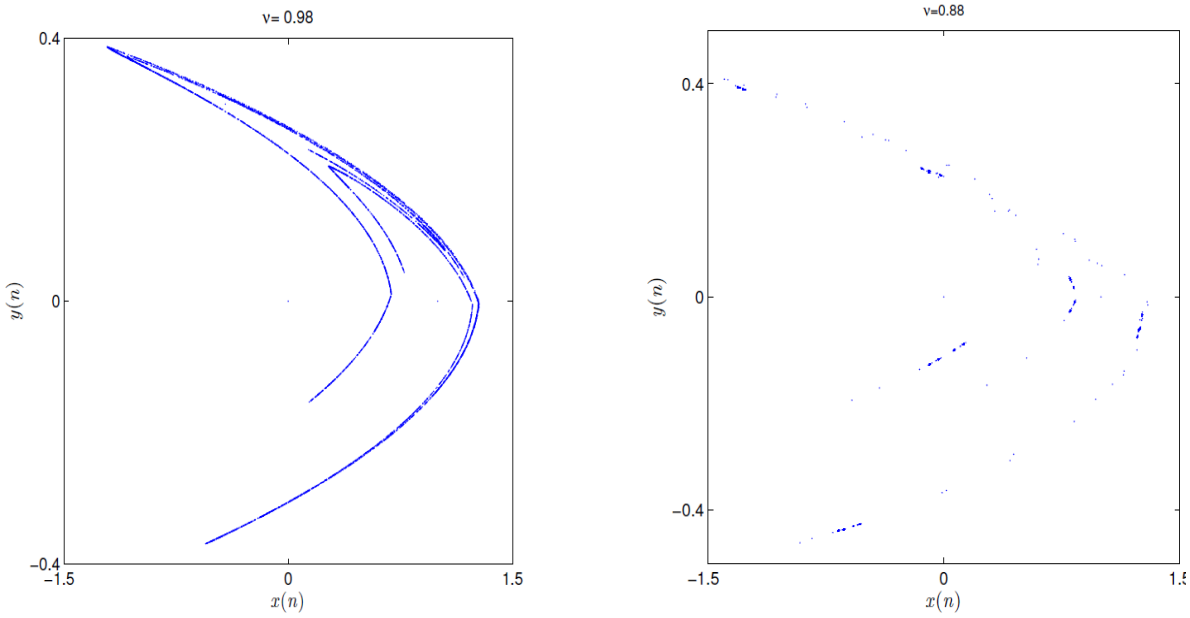


Figure 5 : Chaotic attractor of the fractionalized map for *a*=0.2 and *v*=0.98, periodic attractor for *a*=0.3 *v*=0.88.

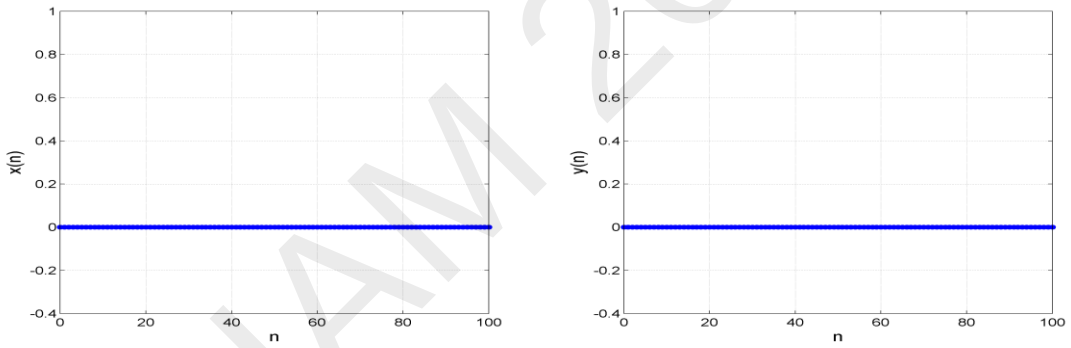


Figure 6: Time evolution of the controlled states of the fractional new map for *a*=0 and α=0.2 and *v*=0.95.

REFERENCES

[1] G.Wu and D.Baleanu, "Discrete fractional logistic map and its chaos". Nonlinear Dyn, vol.75, 2013, pp. 283–287.

[2] G.-C. Wu and D. Baleanu, "Discrete chaos in fractional sine and standard maps". Physics Letters A, 2014, pp.484-487

[3] T.Hu, "Discrete Chaos in Fractional Henon Map", Appl. Math. vol.5, 2014, pp. 2243–2248

[4] E. Zeraoulia, and J.C.Sprott,. ''A unified piecewise smooth chaotic mapping that contains the Hénon and the Lozi systems". Ann. Rev. Chaos Theory Bifurcations Dyn. Syst., vol.1,2012, pp.50–60

Segmentation of Algerian baccalaureate transcripts

Abderrahmane Kefali[1,2]
[1]*Département d'Informatique, Université 8 Mai 1945- Guelma, BP 401, Guelma 24000, Algeria*
[2]*LabGED Laboratory, Badji Mokhtar University, BP 12, 23000 Annaba, Algeria*
kefali.abderrahmane@univ-guelma.dz

Ahlem Obeizi[1]
[1]*Département d'Informatique, Université 8 Mai 1945- Guelma, BP 401, Guelma 24000, Algeria*
ahlemobeizi@yahoo.com

Chokri Ferkous[1,3]
[1]*Département d'Informatique, Université 8 Mai 1945 Guelma, BP 401, Guelma 24000, Algeria*
[3]*Laboratoire des Sciences et Technologie de l'Information et de la Communication (LabSTIC), 8 May 1945 University, Guelma, Algeria.*
ferkous.chokri@univ-guelma.dz

***Abstract*— The digitization and dematerialization of archives is a current trend of a large number of administrations. The simple digitization is not enough; it must be accompanied by techniques facilitating their automatic analysis. This work fits into this context. We propose in this paper a hybrid segmentation approach of Algerian baccalaureate transcripts. It begins with the application of several preprocessings to improve the quality of the input document. Afterwards, RLSA algorithm is used to separate the border of the transcript because it does not matter. Then, a first detection of text lines is performed using RLSA and the detected lines are grouped into blocks. The textual blocks are then identified based on the analysis of projection profiles and their lines are extracted. After that, non-textual blocks are distinguished into tables or graphics using Radon transform. Finally the tables are segmented into columns and cells and their contents are extracted.**

Keywords—segmentation, document image, document analysis and recognition, physical and logical structure.

I. INTRODUCTION

Nowadays, most of the information is still recorded, stored and distributed in paper format. The widespread use of computers for editing documents, with the introduction of Personal Computers and word processing software, and in front of the evolution of computing and the large amount of information, paper document does not remain the primary support. However, electronic document has become an inevitable vector for the exchange of information during a communication process between or outside organizations.

The large number of existing documents and the production of new documents each year raise important questions in the search for effective processing and storage of these documents and the information they contain. This has led to the appearance of new areas of research such as the automatic analysis and understanding of documents, and the recognition of the elements they contain: images, words, characters, manuscript blocks, etc. These elements are organized into structures which carry information about the document content to simplify the reading and interpreting step. In fact, the document has two types of structures: the *physical structure* which describes the layout of the document, the different text blocks, their arrangement relative to each other, etc. and the *logical structure* which is a designation to the semantic content of the document and thus the correspondence between the physical regions and their function. The structure of the document has several important roles other than reading because it also carries information about the content of the document. The latter is responsible for translating the function of the text and the intention of the author at the same time. Hence, understanding a document requires the recognition of its structure in addition to its textual content.

The digitization and dematerialization of archives is a current trend of a large number of administrations. This digitization allows to preserve the records of students, employees, etc., the invoices, purchase orders, mission orders, etc., in an electronic form but this digitization is not enough. It must be accompanied by methods and techniques that facilitating their automatic analysis and search. The present work fits into this context. We are interested in analyzing the images of Algerian baccalaureate transcripts, which is one of the most important documents in the student's file. And as we said previously, the analysis of a document is only possible through the recognition of its structures; we propose in this paper a system aiming to the segmentation or the extraction of the physical structure of the Algerian baccalaureate transcripts.

The remaining of this paper is organized as follows. We first present an overview of the document segmentation approaches existing in the literature. In a second phase, we describe the characteristics of Algerian baccalaureate transcripts. Then, we present our proposed approach for the segmentation of the Algerian baccalaureate transcripts while detailing the various steps included, before concluding.

II. BACKGROUND

A large number of methods have been proposed in the literature for the segmentation or the extraction of physical structure of document images. Most research considers that the segmentation methods may be divided into three classes: bottom-up, top-down, and hybrid [8, 16].

A. Bottom-up approach

The bottom-up approach is guided by data. The methods of this approach start with the lowest level until reaching the highest level in the document page. That means that they start at the connected components level, merge them into words, and then merge these words into lines, the lines in blocks, until the page is completely reconstituted. In this approach, several aspects have been exploited and various techniques and algorithms have been proposed. Without doubt, RLSA algorithm (Run Length Smoothing Algorithm) [23] is one of

the most popular bottom-up methods. This later has been the basis of several other techniques such as [19, 21]. Moreover, we find methods using the connected components, such as [7, 9], methods using window-based filtering ([14]), methods using Voronoi diagrams ([13]), etc.

B. *Top-down approach*

The top-down approach is often used for documents with a well-defined structure. The segmentation in this approach is based on a strong prior knowledge of the document model to cut it into increasingly fine blocks. These methods start with the highest level (the entire image) until reaching the lowest level (connected components). In this approach, several algorithms may be derived. An example of an algorithm using the top-down strategy is the famous X-Y cut algorithm [17], which is more suitable for structures of Manhattan type. Several other techniques have been proposed based on X-Y cut algorithm, such as [12, 18]. Other aspects were exploited. We find methods using the analysis of image background ([1, 22]), methods based on Split and merge ([11]), etc.

C. *Hybrid approach*

The hybrid approach combines bottom-up analysis to extract local primitives and top-down analysis to search for global primitives. Among hybrid techniques we cite [2, 3, 15].

III. CHARACTERISTICS OF ALGERIAN BACCALAUREATE TRANSCRIPTS

After the physical analysis of the various baccalaureate transcripts, distributed over different years (from 1997 to 2015), we noticed that almost every year the format of the transcripts changes (see Fig. 1), but the data remain the same {Frame, Heading, Registration number (R.N) of the student, Student Information, Year, Branch of study, Scores table, ...}. However, the variations are at several levels; for example, at the level of paper quality (standard or special paper), the writing font, the language with which the student's information is written (Arabic or French), the stamp and the signature, the text and background colors,…etc.

According to Fig. 1, it can be noted that the scores table is usually in the middle of the baccalaureate transcript and the average table is below it. It should also be noted that the frame of these two tables is a simple rectangle or a rectangle with rounded corners. Then there are several branches of study in high schools in Algeria, and the branches are different from each other by the number and the content of courses.

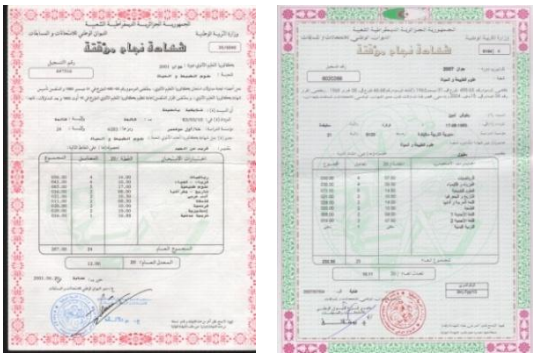
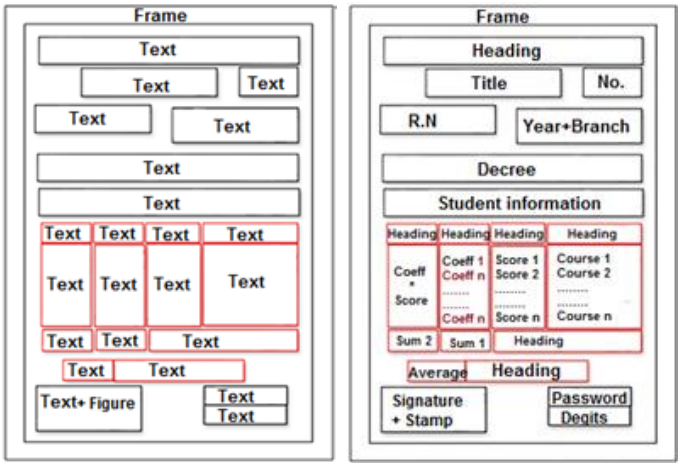


Fig. 1. Examples of baccalaureate transcripts of various formats.



a) Physical structure b) Logical structure

Fig 2. Different levels of structure of an Algerian baccalaureate transcript.

From this physical analysis, we extract the physical and logical structures of the Algerian baccalaureate transcript. They are shown in Fig. 2.

IV. PROPOSED SYSTEM

We describe in this section the proposed approach for the segmentation of Algerian baccalaureate transcripts. The proposed approach consists of several processing steps grouped into two main modules: preprocessing and physical structure extraction, as shown in Fig. 3.

A. *Preprocessing*

Preprocessing gathers a set of operations aimed to eliminate the noise, reduce the degradations, and preserve only

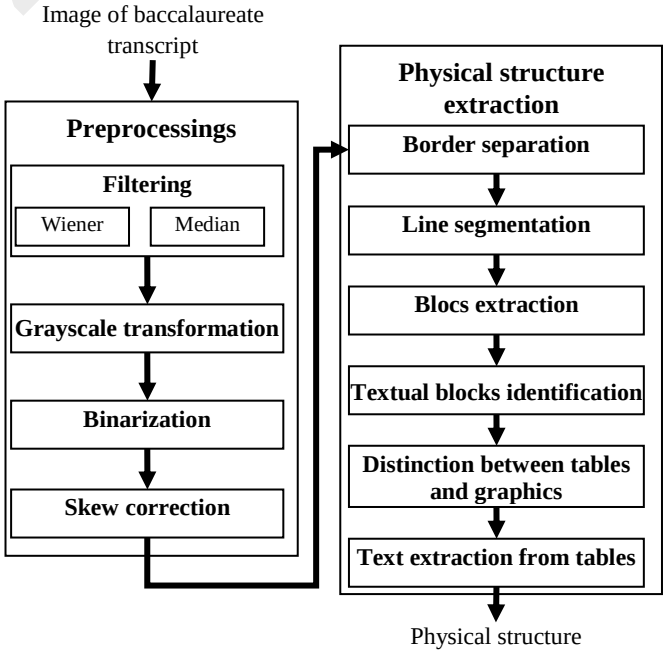


Fig 3. Block diagram of the proposed approach

the useful information of the image. This will prepare the ground for the next steps of processing. In our approach, preprocessing includes filtering, grayscale transformation, binarization, and skew correction.

1) Filtering: To improve the quality of a document image, several filtering may be applied. In our case, the purpose of the filtering is to make better, the details in the transcript image, in order to have good foreground/ background separation results. For this, we use two filters which are very effective for documents preprocessing:

a) Wiener filter: this is one of the most used filters for restoring document images [10]. This filter is effective for processing images whose small details are not enough present. Wiener filter increases the contrast between the texture and the background while smoothing the background.

b) Median filter: Although Wiener filter improves the quality of the document; it does not behave well when the image is strongly noisy. For this, we propose to apply another filter namely the median filter. This filter eliminates impulsive noise where the pixels become randomly scattered on the image surface by generating parasites.

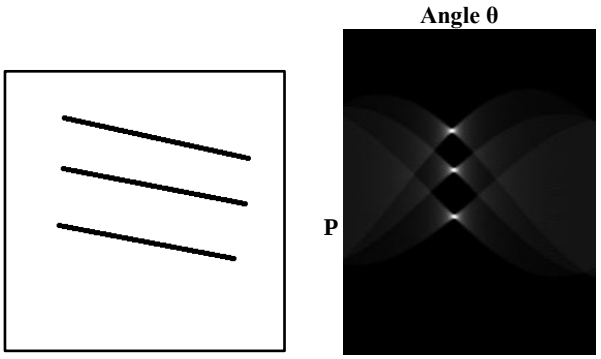
2) Grayscale transformation: this transformation is done simply by replacing the color of each pixel of the image by the average of its values of red, green, and blue.

3) Binarization: it allows to separate the foreground from the background of the image which produces two classes of pixels: background (in white) and scene (in black). In fact, a large number of binarization techniques have been proposed. In our system we chosen to use the method of Sari et al. described in [20] which is a hybrid thresholding method producing, according to the authors, good results for images of degraded documents. This technique runs in two passes. In the first pass, a global thresholding is applied to the image in order to classify the maximum of its pixels (whose gray level is between two global thresholds *T1* and *T2*) into foreground or background. In the second pass, the remaining pixels are assigned to one of two classes: foreground or background based on local analysis

4) Skew correction: the techniques that we will use for segmentation of the baccalaureate transcripts are sensible to the inclination, and because some of our documents are inclined, a step of skew correction is required. However, we used a simple skew correction method based on Radon transformation [5]. The choice of Radon transform is justified by its ability to describe the orientation of the straight lines (which are present in the baccalaureate transcripts and form tables), the simplicity of its implementation, and its independence from the pre-settings operations [4]. Radon transform is a tool allowing to plot the projection histogram of the pixels according to well-defined orientations. According to [6], Radon transform is defined by the following equation:

$$f(p,\theta) = \int_{-\infty}^{+\infty} f(p.\cos(\theta) - s.\sin(\theta), p.\sin(\theta) + s.\cos(\theta))d_s$$

Where θ is the projection angle, p is the coordinate of the point



(a) (b)

Fig. 4. Radon transform of an image, (a) image containing straight lignes, (b) Corresponding Radon space

P on the pixels projection hyperplane and *s* is the coordinate of the point *P* along the perpendicular to this hyperplane.

Thus, Radon transform allows to concentrate the sum of pixels intensities of a straight line at a point of the transformed space. A straight line in an image is then transformed into a point of high intensity in the Radon space, as shown in Fig. 4.

In our proposal, the inclination angle of the document is found by applying Radon transform to the part of the image which may contain straight lines, which is the middle of the image. From the obtained Radon space, the inclination angle θ may be extracted easily and all that remains is to rotate the document image of angle θ.

B. Physical structure extraction

The physical structure of the transcript is organized, as shown in Fig. 2, in blocks. There are textual blocks, composed of one or more text lines, corresponding to the student information, title, etc., and non-textual blocks. The non-textual blocks can be tables (of notes and of average), or graphics (stamp and signature, ...).

To extract the physical structure of the baccalaureate transcripts, we precede in our approach a hybrid segmentation. First, the border of the transcript is separated from the image using a bottom-up technique because it carries no relevant information. Then, the lines and the blocks are extracted from the image of transcript without border. After that, the textual blocks are identified and the other blocks are localized. Finally, the localized tables are segmented and their information is extracted.

1) Border separation: according to the physical study of the baccalaureate transcripts that we conducted, we noted that the most of transcripts contains different formats of the frame surrounding the document information: frame in the form of a rectangle and so it is formed of a single connected component, frame in the form of a series of stars or other geometric shapes, etc. There are also other transcripts that do not contain any borders.

To separate the border of the baccalaureate transcript, we proceed with a method based on RLSA algorithm. The idea is to gather the pixels of the frame or the border into a single unit, and to separate them from the document, the task that may be accomplished efficiently using RLSA algorithm. This

latter allows to connect black pixels separated by less than *n* white pixels according to horizontal or vertical direction. In fact, the border takes the form of a rectangle formed of four sides (top, bottom, right, left). Separating the border therefore requires the localization of its four sides on the document. Thus, the physical study that we did allows us to know the approximate location of the four sides of the border. The latter may differ slightly from one document to another, but they are always located in the first parts (top, bottom, right, and left) of the image with a thickness that never exceeds the value (document width / 10). To find the horizontal sides (top and bottom) of the border, for example, we apply the RLSA algorithm horizontally on the top and bottom parts of the image with a threshold *n* = (image width * 10%). The localization of the two vertical sides (right and left) is done in the same way but by applying a vertical RLSA on the right and left parts of the image, with a threshold *n* = (image width * 20%). The values of the threshold *n* have been chosen by experimentation so that they allow to connect the closest components of the border.

The application of RLSA algorithm on the parts of the image containing the horizontal (or vertical) sides of the frame leads to connect the black pixels of the border that are near along the horizontal (or vertical) direction. At the end the border becomes composed of a single object (Fig. 5). A connected components labeling is then performed only on the parts of image containing the four sides of the border, in order to gather all the pixels composing the border into a single unit (connected component). This connected component representing the frame is finally separated from the document image because its presence may influence the results of the following segmentation steps.

2) Lines segmentation: Once the border is separated from the image, only the relevant elements of the transcript (text, tables, ...) remain in black on a white background. And since the image is well oriented (after the skew correction step), the text lines of the transcript may be extracted by a RLSA smoothing. Thus a horizontal RLSA smoothing is applied to the resulting image of the preceding steps in order to eliminate the spaces between the words close of the same text line. In addition, a vertical RLSA smoothing is applied in order to connect the diacritic marks to their corresponding words (because the transcripts are in Arabic and the diacritics marks are strongly present in Arabic script).

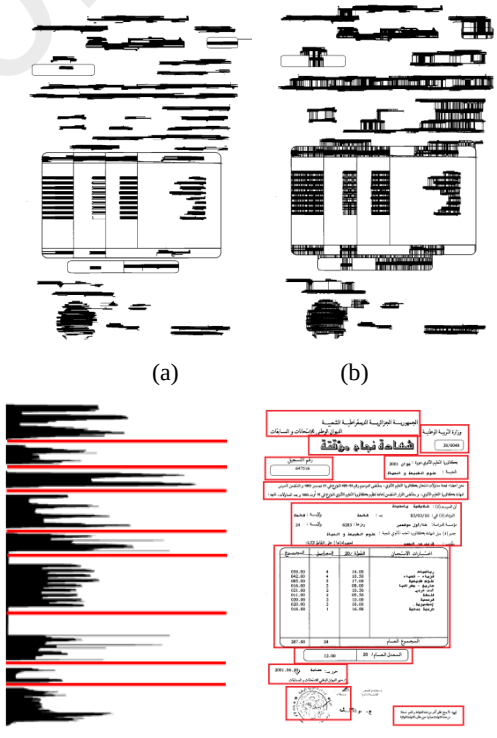
(a) (b)

Fig. 5. Border detection, (a) binarized transcript image, (b) border detected using RLSA algorithm.

Noting that the transcript contains several blocks (heading, title, decree, student information, scores table,...) and that two blocks may overlap horizontally. For example, from Fig. 2, the block "R.N" and the block "Year + Branch" overlap horizontally. Therefore, the horizontal RLSA threshold must be chosen so that it connects the words of a text line of a block, and at the same time does not allow to connect the text lines which belong to two horizontally overlapping blocks. Fig. 6.a shows the result of this step.

3) Blocks extraction: the blocks can be extracted using RLSA algorithm together with the projection profile analysis. At first, we apply RLSA technique vertically with a predetermined threshold, on the resulting image of the previous RLSA smoothing (Fig 6.b). This second smoothing RLSA aims to connect the text lines of the same textual block or to connect the pixels, vertically close, of a non-textual block. From the remark that the interline distance is smaller than the distance between successive blocks, the smoothing threshold must be chosen carefully. It must be large enough to allow the connection of the lines of the same block and at the same time must be insufficient to connect the blocks together.

In a second time, from the smoothed image by RLSA, we proceed to the segmentation of the document into blocks based on the analysis of horizontal and vertical projection profiles. The histogram of horizontal projections is first obtained by calculating the number of black pixels in each line of the image resulting from RLSA smoothing. As the document is well aligned, the corresponding histogram of horizontal



(a) (b)

(c) (d)

Fig. 6. Line and blocks extraction, (a) lines detected using RLSA, (b) first extraction of blocks using RLSA, (c) projection profiles of the smoothed image, the valleys are in red, (d) blocks extracted.

projections will consist of peaks and valleys, representing the blocks and spaces between them respectively (Fig. 6.c). The space between two consecutive valleys describes the height of a block (or horizontally overlapped blocks). Then, we proceed to a second analysis of vertical projection profiles on each block resulting from the first analysis. This second analysis is performed to refine the block segmentation, by separating the blocks that overlap horizontally and that are considered as a single block during the first analysis. Thus, if the histogram has several peaks and valleys, this indicates that several horizontally adjacent blocks are present and to separate them it is enough to segment in the valleys. The final result of blocks extraction is illustrated in Fig. 6.d.

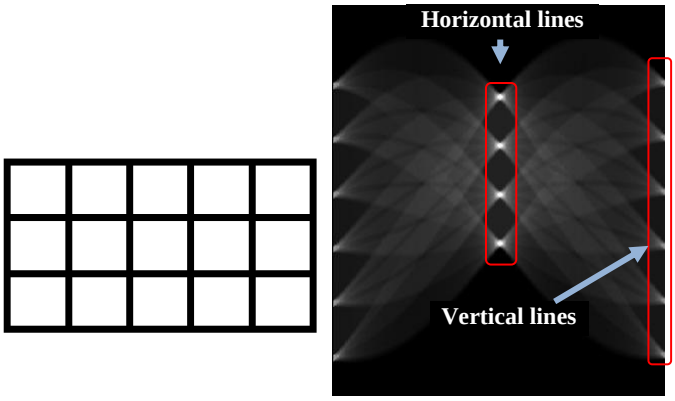
4) Textual blocks identification and lines extraction: As we said before, the blocks may be textual, formed of several text lines, or non-textual. To identify the textual blocks among all the extracted blocks, we examine each block and test whether it consists of several text lines or not. In fact, we have already extracted the text lines using RLSA during the first step of segmentation, but this first segmentation may not be precise because of the presence of noise pasting two lines together, for example. This is why we propose to carry out a second extraction of text lines to confirm.

However, the presence of text lines in a block may be tested based also on the analysis of projection profiles. Thus, we calculate the histogram of horizontal projections of each block extracted from the binarized image (before segmentation). The presence of several global minima, with few black pixels, in the histogram may show inter-line spaces for a textual block. A text line will be between two successive minima. To ensure that this is really a textual block, we apply in a second time the analysis of vertical projection profile for each text line obtained from the first analysis. The presence of several white valleys in the histogram shows certainly interword spaces. Obviously, the absence of interline and interword spaces confirms that it is not a textual block.

5) Distinction between tables and graphics: Once the textual blocks are identified, only non-textual blocks remain. These can be tables or graphics (logos, stamp and signature). To distinguish tables from graphics we rely again on the use of Radon transform. In fact, Radon transform seems a good choice because it allows to detect the presence of straight lines. These are considered as the key element to discriminate the presence of tables. Thus, Radon transform is applied to each non-textual block extracted from the binarized image (before segmentation) in order to check for the presence of straight lines in this block.

As we have already said, Radon transform returns peaks in the form of dots, which signal the presence of straight lines. Fig.7 shows the result of Radon transform on a table image. The Radon space thus constructed indicates the presence of 10 points of strong luminosity indicating the presence of 10 lines in the table image: 4 points corresponding to the horizontal lines and 6 points corresponding to the vertical lines.

In the case where the non-textual block does not contain straight lines (graphic), Radon transform of this block does not present any peak as a point of strong intensity (Fig. 8).



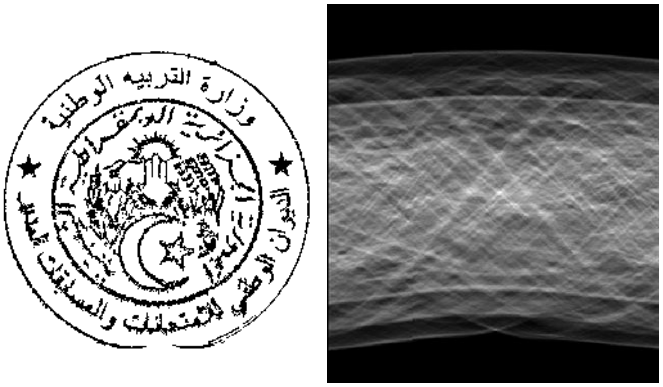
(a) (b)

Fig. 7. Detection of a table, (a) table image, (b) its Radon transforme

6) Text extraction from tables: the last step in our approach is the extraction of text from the tables identified in the previous step. To detach the text from a table, we proceed to the segmentation of this table, extracted from the binary image, in columns and then in cells, and we also make use of the analysis of projection profiles.

a) Segmentation into columns: The segmentation of a table into columns is done by analyzing the profiles of vertical projections. But before that, it is necessary to remove the table's border. To do this, one has only to perform a connected components labeling, only on the area of the image containing the table. The largest connected component is the border of the table and it will be eliminated. Then, the histogram of vertical projections is built on the area of the table without border. The sequences of white pixels of the histogram correspond certainly to inter-column spaces. The separation of columns is done by segmenting the table in these sequences.

b) Columns segmentation into cells and extraction of their content: The cells are segmented from the columns by analysis of horizontal projection profiles. Thus, on each column obtained, the histogram of horizontal projections is calculated. The peaks of the histogram represent the cells of the column and the valleys correspond to the spaces between the cells. The separation of the cells and the extraction of their contents is done by dividing the column in these valleys.



(a) (b)

Fig. 8. Detection of graphic, (a) stamp image, (b) its Radon transform

V. CONCLUSION

In this paper, we proposed a segmentation approach for a particular type of document, namely Algerian baccalaureate transcripts. These are of considerable importance in the student's file. The goal is to develop the core of a system for digitization, analysis, recognition, and retrieval of archives documents within Algerian universities.

The proposed approach may be classed within the category of hybrid methods and it consists of several stages of processing gathered in two modules. A first module incorporating various preprocessings, namely filtering, grayscale transformation, binarization, and skew correction, aimed to improve the quality of the input images and preparing them for the following steps. The second module aims to extract the physical structure of the baccalaureate transcripts by applying several steps. First, the border of the transcript is removed because it does not provide any relevant information. After that, a first lines segmentation is performed based on the RLSA algorithm. The lines are then combined into blocks by applying RLSA again and then by analyzing projection profiles. The textual blocks are next identified and their lines are extracted by alternating horizontal and vertical projection profiles. Non-textual blocks are separated into tables or graphics using Radon transform. Finally the tables are segmented into columns and cells and their contents are extracted.

At this stage, the system is under construction. The first tests on the steps implemented until now have shown good results, particularly in terms of pre-processing, border separation, and blocks extraction, which supports the choices made. Once the implementation of the proposed approach is completed, the experiments performed, the results obtained, the advantages and disadvantages will be discussed in another separate article.

REFERENCES

[1] A. Antonacopoulos, "Page segmentation using the description of the background," Computer Vision and Image Understanding, vol. 70, No. 3, pp. 350-369, 1998.

[2] A.S. Azokly, "Une approche uniforme pour la reconnaissance de la structure physique de documents composites fondée sur l'analyse des espaces," doctorat dessertation, Université de Fribourg-France, 1995.

[3] H.S. Baird, E.J. Susan, and J.F. Steven, "Image segmentation by shape-directed covers," Proc. 10th International Conference on Pattern Recognition, vol. 1, pp. 820-825, 1990.

[4] A. Ben Salah, "Maîtrise de la qualité des transcriptions numériques dans les projets de numérisation de masse," doctoral dissertation, Université de Rouen-France, 2014.

[5] R.N. Bracewell, "Two-Dimensional Imaging," Englewood Cliffs: Prentice Hall, vol. 247, 1995, pp. 505-537.

[6] P. Courmontagne, "Transformée de radon et filtrage : Application à la détection de sillages de mobiles marins," TS. Traitement du signal, vol. 15, No. 4, pp. 297– 307, 1998.

[7] D. Drivas, and A. Amin, "Page Segmentation and Classification Utilizing Bottom-Up Approach," Proc. 3rd International Conference on Document Analysis and Recognition, pp. 610-614, Montreal, Canada, 1995.

[8] S. Eskenazi, P. Gomez-Krämer, and J.M. Ogier, "A comprehensive survey of mostly textual document segmentation algorithms since 2008," Pattern Recognition, vol. 64, pp. 1-14, 2017.

[9] J. Fisher, S. Hinds, and K. D'Amato, "A Rule-Based System for Document Image Segmentation," Proc. 10th International Conference on Pattern Recognition, pp. 113-122, Atlantic City, USA, 1990.

[10] B. Gatos, I. Pratikakis, and S.J. Perantonis, "Adaptive degraded document image binarization," Pattern recognition, vol. 39, No. 3, pp. 317-327, 2006.

[11] S.L. Horowitz, and T. Pavlidis, (1972). "Picture segmentation by a traversal algorithm," Comput. Graphics Image Process, vol. 1, pp. 360–372, 1972.

[12] A.K. Jain, and Y. Bin Yu, "Document representation and its application to page decomposition," Pattern Analysis and Machine Intelligence, vol. 20, No.3, pp. 294-308, 1998.

[13] K. Kise, A. Sato, and M. Iwata, "Segmentation of Page Images Using the Area Voronoi Diagram," Computer Vision and Image Understanding, vol. 70, No. 3, pp. 370-382, 1998.

[14] F. Lebourgeois, Z. Bublinski, and H. Emptoz, "A Fast and Efficient Method for Extracting Text Paragraphs and Graphics From Unconstrained Documents," Proc. 11th International Conference on Pattern Recognition, pp. 272-276, The Hague, 1992.

[15] J. Liu, Y. Tang, Q. He, and C. Suen, "Adaptive Document Segmentation and Geometric Relation Labeling: Algorithms and Experimental Results," Proc. 13th International Conference on Pattern Recognition, pp.763-767, Vienna, Austria, 1996.

[16] S. Mao, A. Rosenfeld, and T. Kanungo, "Document structure analysis algorithms: a literature survey," In Document Recognition and Retrieval X, vol. 5010, International Society for Optics and Photonics, 2003 pp. 197-208.

[17] G. Nagy, and S. Seth, "Hierarchical representation of optically scanned documents," Proc. 7th International Conference on Pattern Recognition (ICPR), pp. 347-349, 1984.

[18] G. Nagy, J. Kanai, M. Krishnamoorthy, M. Thomas, and M. Viswanathan, "Two Complementary Techniques for Digitized Document Analysis," Proc. ACM Conference on Document Processing Systems, pp. 169-176, Santa Fe, New Mexico, USA, December 1988.

[19] N. Nikolaou, M. Makridis, B. Gatos, Nikolaos Stamatopoulos, and Nikos Papamarkos, "Segmentation of historical machine-printed documents using adaptive run length smoothing and skeleton segmentation paths," Image and Vision Computing, vol. 28, No. 12, pp. 590–604, 2010.

[20] T. Sari, A. Kefali, and H. Bahi, "Text extraction from historical document images by the combination of several thresholding techniques," Advances in Multimedia, vol. 2014, pp. 11, 2014.

[21] Z. Shi, and V. Govindaraju, "Line separation for complex document images using fuzzy runlength," Proc. International Workshop on Document Image Analysis for Libraries, pp. 23–24, 2004.

[22] A.L. Spitz, "Recognition processing for multilingual documents," Proc. International Conference on Electronic Publishing, Document Manipulation and Typography, pp. 193-205, Gaithe rsburg, Maryland, 1990.

[23] F. Wahl, K. Wong , and R. Casey, "Block Segmentation and Text Extraction in Mixed Text/Image Documents," Computer Vision Graphics, and Image Processing, vol. 20, pp. 375-390, 1982.

An efficient Classification of Epileptic seizures using autoregressive model Based on Deep Learning and SVM

Abdelouahab ATTIA
Computer Science Department - Faculty of Mathematics and informatics Mohamed El Bachir El Ibrahimi University of Bordj Bou Arreridj,
B.B.A 34000, Algeria
attia.abdelouahab@gmail.com

Abdelouahab MOUSSAOUI
Computer Science Department - Faculty of Sciences
Ferhat Abbas University - Setif I ,
Setif 19000, Algeria
moussaoui.abdel@gmail.com

Youssef CHAHIR
GREYC Laboratory - CNRS UMR 6072
Departement of Computer Science
University of Caen Lower-Normandy-France
youssef.chahir@unicaen.fr

Mourad CHAA
ELEC Laboratory – Faculty of new technology of information and communication Ouargla university,
Ouargla 30 000, Algeria
chaa500@yahoo.com

***Abstract*—The classification and detection of epileptic seizure with machine learning methods have become a major key in the diagnosis of epilepsy. Generally, EEG signals are non-stationary and complex. Given this, it has been difficult for classical methods to extract practical information. To encounter this problem, the current paper introduces a new framework for the classification of an epileptic seizure using autoregressive model (AR) and deep learning (DL) autoencoder with Support Vector Machine as a classifier (SVM). First, features' extractions have been accomplished by Yule-Walker (Y-W) and Burg methods. Then, deep learning has been used for improving the features given by such algorithms. These vectors have been combined in one single vector. Then, SVM classifier has been used for the training and classification stage. To validate the performance of the proposed framework, the Bonn database has been employed. In the experimental stage, two types of classification have been studied (healthy non-epileptic to Interictal or Ictal). The obtained results with an average accuracy are (94.50 %,100.00%). They have clearly demonstrated the efficiency of the proposed framework.**

Keywords— epileptic seizures; sparse Autoencoder; Autoregressive model; SVM classifier;

I. INTRODUCTION

The electroencephalogram (EEG) has become a powerful clinical tool in the investigation and diagnosis of neurological disorders such as epilepsy [1]. Basically, EEG signals are non-stationary process. They contain a large range of frequency components that is between 0.5 and 30 Hz overlapping with each other [2]. Thus, the process of EEG evaluation, analysis and interpretation has been a complex and difficult task with no explicit criterion [3,4]. Hence, traditional recognition of epilepsy has been incompetent for providing doctors with more information. To overcome this limitation, automatic detection of epileptic seizures has become an alternative solution that can precisely and efficiently highlight undetectable information to better diagnosing epilepsy. In the context of automatic detection and classification epileptic seizures, several works exploring EEG data have been introduced; including , Alkan et al [1] Employed Logistic Regression (LR) and Multilayer Perceptron Neural Network (MLPNN) classifier, which achieve a classification accuracy of 92.00 %. Yalcin et al [4] have analyzed Epilepsy using artificial neural network learned by Particle Swarm Optimization. PSO have provided an overall accuracy of 99.67 %. Chandaka et al. [5] have used Cross-correlation with SVM classifier and the final accuracy was 95.96 %. Guo et al. [6] have introduced a framework based on line length feature and artificial neural networks for detecting automatic epileptic seizure that have achieved a better accuracy of 99.60 %. Also, the same authors have employed the genetic programming and k-nearest neighbor KNN classifier [7] and got an accuracy of 99.20 % . Nicolaou et al.[8] have proposed a model based on the permutation entropy and SVM to detect epileptic seizures. They have reported an accuracy of 93.55 %. Subasi and Gursoy [9] have employed the dimensionality reduction methods such as Linear Discriminant Analysis (LDA), Principle Component Analysis (PCA), and Independent Component Analysis (ICA) with SVM for EEG signal classification. Thus, they have provided comparative studies and achieved a better accuracy of 98.75% 99.50% and 100.00% for (PCA with SVM), (ICA with SVM)and (LDA with SVM) respectively. Lin et al.[10] have introduced a new framework for Classifying Epileptic EEG data by employing the stacked sparse autoencoder (SSAE) and a softmax classifier. The authors have reported 100% classification accuracy. Furthermore, the choice of a method of feature extraction plays an important role in the classification performance. To address this problem, the current paper introduces a new framework for automatic detection and classification of epileptic seizures by using a parametric method known as autoregressive model (AR) [11] and improved by sparse autoencoder (SAE) based on deep learning. However, AR model has been used in features extraction of EEG signals [12]. This permits selecting coefficients which can be used directly by the classifier algorithms. In addition, AR modeling is a very useful tool for the signals which have a high

frequency resolution and the smoother spectra [13]. In this study, two kinds of AR modeling (i) Yule-Walker algorithm and (ii) Burg algorithm have been used in order to estimate the model coefficients of the input EEG signal [14, 15]. The obtained coefficients have been improved by SAE. Then, SVM classifier has been applied. Experimental results have clearly proven that the proposed framework outperforms the traditional methods.

The remaining paper is organized as follows: Section 2 describes the proposed method including the employed tools and data. Section 3 presents the conducted experiments and results. At last, a conclusion is given.

II. THE PROPOSED METHOD

The proposed method for epileptic seizure classification is made up of four principal stages. The first stage consists in extracting features using the AR Yule-Walker method (AR-Y-W), extracting features using (AR-Burg) and combining features vectors of the Yule-Walker and AR Burg vectors. After that, deep learning has been applied to features vectors to improve values of features. Then, SVM method has been employed in the classification stage. The final stage is the evaluation performances of the classifier. Fig. 1 clearly illustrates the flowchart of the introduced method.

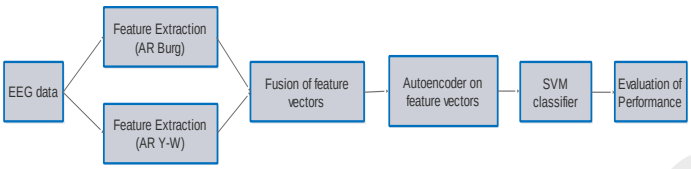


Fig. 1. Diagram of the proposed system

A. *Feature extraction using autoregressive coefficients*

In this work, parametric methods have been employed for feature extraction to provide the parameters representing EEG signals denoted as feature vectors that can be used for the classification stage. Particularly, the focus has been placed on the use of an autoregressive method (AR). Two kinds of AR model styles have been employed: (i) Yule-Walker algorithm and (ii) Burg algorithm. This will be described in the following sections assuming that the input signal EEG X(n) n varied n=1,2, 3 …., N where N is the length of EEG signal.

1) Autoregressive Modelling for EEG

AR model has become one of the important parametric methods that have been applied in many studies to model EEG signals. AR model permits describing the EEG signal as a linear representation. The regression model of the EEG signal has been provided by the following formula [16]:

$$X(n) = -\sum_{k=1}^{p} a_k\, x(n-k) + e(n) \qquad (1)$$

Where a_k denotes the AR coefficient and p represents the given model order give. e(n) denotes the error term independent of the previous samples that have been assumed to be white Gaussian noise with zero mean and variance σ^2. As noted above, the AR parameters have been estimated by using both yule-walker and burg methods described in what follows:

2) Yule-walker.

The Yule-Walker algorithm (Y-W) is based on computing a least-squares fit of the predicts model to estimate the model coefficients of the input EEG signal. However, the Y-W algorithm performs by using the equations introduced by Yule-Walker [14, 15]:

$$E = \sum_{n=1}^{N} (e(n))^2 = \sum_{n=1}^{N} \left(x(n) - \hat{x}(n) \right)^2 \qquad (2)$$

$$e(n) = x(n) - \hat{x}(n) \qquad (3)$$

Where $\hat{x}(n)$ is the predicted value.

$$E = \sum_{n=1}^{N} \left(x(n) - \sum_{n=1}^{N} a_k\, x(n-k) \right)^2 \qquad (4)$$

$a_k$is predicted to minimize error e(n). Mean square value of the error will be minimum if $\frac{\partial E}{\partial a_k} = 0$

We obtain:

$$\sum_{k=1}^{P} a_k R(k-i) = r(i) \qquad (5)$$

Where r (i), i = 0, 1, …, p - 1 denotes the autocorrelation coefficients for the lag i and can be formed into the following matrix expression

$$\begin{bmatrix} r(0) & r(1) & \cdots & r(p-1) \\ r(1) & r(0) & \cdots & r(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ r(p-1) & r(p-2) & \cdots & r(0) \end{bmatrix} \times \begin{matrix} a(1) \\ a(2) \\ \vdots \\ a(p) \end{matrix} = \begin{matrix} r(1) \\ r(2) \\ \vdots \\ r(p) \end{matrix} = Ra = r \qquad (6)$$

Finally, these equations are the estimated Yule-Walker equations and this is the autocorrelation method of linear prediction.

$$\alpha = R^{-1} r \qquad (7)$$

3) Burg Algorithm

The Burg Algorithm is a common method. It has been used to estimate the parameters of an AR model because it is different for other methods that guarantee to generate a stable model. The algorithm is a recursive method based on the lattice filter structure in order to minimize the forward and backward prediction error [17]. The algorithm is as follows:

- **Step1**: Calculate the initial values of error variance, Forward error and Backward error by the given equations respectively

$$\sigma^2(0) = \frac{1}{N}\sum_{0}^{N-1}(x(n))^2 \quad (8)$$

$$e_n(0) = x(n) \quad (9)$$

$$b_{n-1}(0) = x(n-1) \quad (10)$$

- **Step2**: Calculate reflection coefficient and error variance by the given equations respectively:

$$\pi_m = \frac{\sum_{n=m}^{N-1} b_{n-1}(m-1)\, e_n(m-1)}{\sum_{n=m}^{N-1}\left(e_n^2(m-1) + b_{n-1}^2(m-1)\right)} \quad (11)$$

$$\sigma^2(m) = \left(1 - |\pi_m|^2\right)\sigma^2(m-1) \quad (12)$$

- **Step3**: Update Error and AR coefficients

AR coefficients:

$$\begin{cases} a_k(m) = a_k(m-1) + \pi_m a_{m-k}(m-1) & \text{if } m > 1 \\ a_m(m) = \pi_1 & \text{if } m = 1 \end{cases} \quad (13)$$

Forward Error:

$$e_n(m) = e_n(m-1) + \pi_m b_{n-1}(m-1) \quad (14)$$

Backward Error :

$$b_n(m) = b_{n-1}(m-1) + \pi_m e_n(m-1) \quad (15)$$

- **Step** 4: Repeat steps 2 and 3 (with m incremented by one) until the selected model order p is reached.

4) The Akaike Information Criterion

The Akaike Information Criterion (AIC) [18] has been employed in the evaluation of the AR model coefficients. However, the AIC method permits selecting an adequate AR model order P [19,20]. Assuming that the input has Gaussian statistics, the AIC for an AR process is defined by:

$$AIC(p) = \ln(\sigma^2) + \frac{2p}{N} \quad (16)$$

Where σ^2 is white noise variance.

B. Improve Feature vectors by sparse autoencoder

The main objective of the current paper is the use of the autoencoder to enhance the features vectors extracted from EEG signals using both AR models Yule walker and Burg. Basically, an autoencoder is an unsupervised learning method that is trained to reconstruct its input at its output (encoding). It denotes an artificial neural network. The autoencoder consists of two essential parts; the encoder and the decoder. Thus, three layers have been employed: (i) an input layer, (ii) an output layer and (iii) one hidden layer to improve the feature vectors extracted by AR models from EEG signals. As depicted in figure 2, these layers linked between them are illustrated.

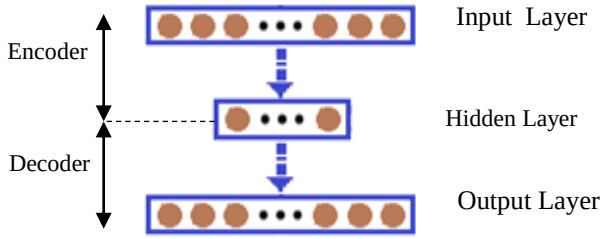


Fig. 2. An autoencoder structure

Let V be a set of training features vectors provided by AR model that the input to an autoencoder. Then, the encoder maps this data to another data Z using the following formulae:

$$z^{(1)} = f^{(1)}\left(w^{(1)}x + b^{(1)}\right) \quad (17)$$

Then, the decoder maps the encoder representation z back into an estimate of the original input matrix,x, as follows:

$$\hat{x} = g^{(2)}\left(w^{(2)}x + b^{(2)}\right) \quad (18)$$

Where $w^{(1)}$ and $w^{(2)}$ are the weights matrix between the input and the hidden layer (encoder) and the hidden and the output layer (decoder) respectively. $b^{(1)}$ and $b^{(2)}$ are bias vectors for the hidden and the output layers respectively. The transfer functions of the encoder and decoder are represented by $f^{(1)}$ and $g^{(2)}$ correspondingly. The Logistic sigmoid activation function has been used for both the encoder and the decoder. It is calculated by:

$$\log sig(v) = \frac{1}{(1+\exp(-v))} \quad (19)$$

Since no labeled data is required, training an autoencoder is unsupervised. It is worth noticing that the training process still relies on the optimization of a cost function. The cost function measures the error between the input $x \in \mathbb{R}^D$ and its reconstruction at the output $\hat{x} \in \mathbb{R}^D$. Training a sparse autoencoder consists of adjusting the mean squared error function as follows:

$$E = \frac{1}{N}\sum_{n=1}^{N}\sum_{m=1}^{M}\left(x_{mn} - \hat{x}_{mn}\right)^2 + \alpha \cdot \lambda_{weight} + \beta \cdot \lambda_{sparsity} \quad (20)$$

Where α and β refer to the coefficients of the regularization term and the sparsity regularization term respectively. The $\lambda_{sparsity}, \lambda_{weights}$ are calculated as follows:

$$\lambda_{weights} = \frac{1}{2}\sum_{h}^{H}\sum_{j}^{n}\sum_{i}^{m}\left(w_{ij}^{(h)}\right)^2 \quad (21)$$

$$\lambda_{sparsity} = \sum_{i=1}^{D} \xi \log\left(\frac{\zeta}{\hat{\zeta}_i}\right) + (1-\zeta)\log\left(\frac{1-\zeta}{1-\hat{\zeta}_i}\right) \qquad (22)$$

where H denotes the number of hidden layers, n refers to the number of observations, and m stands for the number of the training data. Then $\hat{\xi}_i$ represents the average output activation measures of a neuron i that is given by:

$$\hat{\zeta}_i = \frac{1}{n}\sum_{j=1}^{n} h\left(w_i^{(1)T} x_j + b_i^{(1)}\right) \qquad (23)$$

Where x_j represents the j the training sample. $w_i^{(1)T}$ and $b_i^{(1)}$ are the i row of the weight matrix and the i entry of the bias vector respectively. The parameters of the sparse autoencoder have been fixed as $\alpha = 0 \cdot 5$, $\beta = 10$ to form the feature vector based on the proposed method.

C. *support vector machine*

The support vector machine (SVM) classifier is a supervised learning method based on the large margin principle in separating data into two groups [9]. The choice of kernel function provides different classifiers that are linear and nonlinear separable [5]. In addition, the parameters of SVM play an important role in the classification accuracy. Thus, the radial basis functions (RBF) have been used as kernel functions of SVM classifier that achieve good results. The RBF is given by:

$$K(x, y) = \exp\left(-\frac{|x-y|^2}{2\sigma^2}\right) \qquad (24)$$

III. Experimental Classification Results

A. *Dataset*

The Bonn database [21] has been commonly used in epilepsy study and seizure detection research. Bonn database was collected by Ralph Andrzejak et al using the standard 10–20 electrode system. The length of each recording is 173.61 × 23.6 ≈ 4097 samples [21] and their description is as follows: The whole data contains five sets (Z, O, N, F, S) where each one of them has 100 single channels of EEG segment with duration equals 23.6 sec. These collections are given from three groups. The first group was collected from five subjects with healthy volunteers and eyes open, and also from the same subject with eyes closed denoted as healthy (nonepileptic) group. The second group is the interictal group. The third group is called the ictal group for the seizure activity. Fig 3 presents an example of each set.

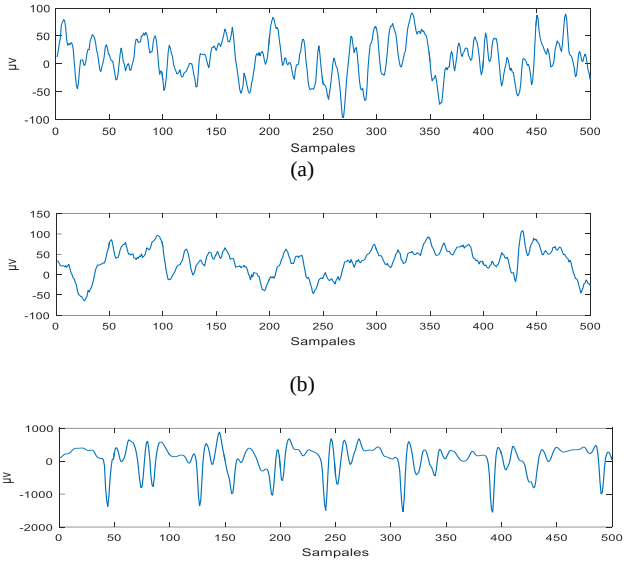


Fig. 3. : Exemplary typical EEG signals of the sets: (a) class (Z, O), (b) class(N, F), and (c) class(S).

B. *3.2 Performance Evaluation Measurements*

The performance metrics have been used are accuracy, sensitivity and specificity measures.

Sensitivity (SEN):

$$SEN = \frac{TP}{(TP + FN)} \qquad (25)$$

Specificity (SEP):

$$SPE = \frac{TN}{(TN + FP)} \qquad (26)$$

Classification accuracy (ACC):

$$ACC = \frac{(TP + TN)}{N} \qquad (27)$$

Where TP is True positive, TN denotes True negative, FP stands for False positive and FN represents False negative and N number of signals.

C. *Discussions*

Two types of experiments have been realized. In the first experiments, the features have been extracted by using Yale-Walker and Burg method with combining the features vectors of each algorithm. In the end, SVM classifier has been applied. In the second experiments, the (DL) has been applied to the AR coefficients for improving these coefficients. After this, SVM classifier has been applied. In these experiments, the data sets have been divided into three classes: (i) healthy (nonepileptic) –Z, O; (ii) Interictal–N, F; and (iii) Ictal–S. Then, every signal has been divided into the onset of segments 23.6s. the notation A, B, C, D, E have been employed for the classes Z, O, N, F, S respectively to make a comparison with other works. Also in the conducted experiments, the best model order P has been selected by using the AIC metric for both Y-W and burg AR model. Finally, the classification has been studied as follows:

(i) healthy no-epileptic (Normal) to Interictal (sets A, B) to (sets C, D). (ii) healthy no-epileptic (Normal) to Ictal (sets A, B) to (set E).

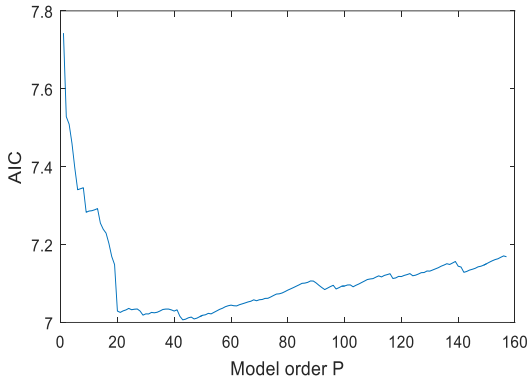


Fig. 4. An illustrative example of the AIC(P) variation

Table 1 illustrates in details the obtained results of the first experiment concerning the first part of the proposed scheme. Table 2 presents the average of the classification evaluation performance. It can be obviously seen from Table 1 that the highest classification accuracies of Sets (A and E)and (B and E) are 99.00 %, 98.00 %, respectively. Then, we achieve an average classification accuracy of 98.50 %, 100.00% and 97.00% for the sensitivity and specificity respectively for healthy no-epileptic (Normal) to Ictal.

TABLE I. RESULTS OF CLASSIFICATION USING (AR W-Y WITH BURG AND SVM))

	healthy no-epileptic (Normal) set A			healthy no-epileptic (Normal) set B		
	To Interictal (sets C, D		To Ictal set E	To Interictal (sets C, D		To Ictal set E
	AC	AD	AE	BC	BD	BE
ACC	0.9600	0.9500	0.9900	0.9400	0.9600	0.9800
SEN	0.9400	0.9800	1.0000	0.9000	0.9600	1.0000
SPE	0.9800	0.9200	0.9800	0.9800	0.9600	0.9600

TABLE II. THE AVERAGE OF PERFORMANCE MEASURE OF THE FIRST EXPERIMENT (AR Y-WAND BURG +SVM)

	healthy no-epileptic (Normal) To Interictal	healthy no-epileptic (Normal) To Ictal
ACC	95.25%	98.50%
SEN	94.50%	100.00%
SPE	96.00%	97.00%

Table 3 depicts in details the results of the second experiment of the second part of the proposed framework. Table 4 presents the average of the classification evaluation performance. It is clearly seen from Table 3 that the highest classification accuracies of Sets (A and E) and (B and E) are 100.00 %, 100.00 %, respectively. Then, an average classification accuracy of 100.00 %, 100.00%, and 100.00% is achieved for the sensitivity and specificity respectively for healthy no-epileptic (Normal) to Ictal

TABLE III. RESULTS OF CLASSIFICATION USING (AR W-Y +BURG IMPROVED BY DL AND SVM))

	healthy no-epileptic (Normal) set A			healthy no-epileptic (Normal) set B		
	To Interictal (sets C, D		To Ictal set E	To Interictal (sets C, D		To Ictal set E
	AC	AD	AE	BC	BD	BE
ACC	0.9600	0.9300	1.0000	0.9500	0.9400	1.0000
SEN	1.0000	0.9100	1.0000	0.9000	0.9200	1.0000
SPE	0.9200	0.9500	1.0000	1.0000	0.9600	1.0000

TABLE IV. THE AVERAGE OF PERFORMANCE MEASURES OF THE SECOND EXPERIMENT (AR-DEEP LEARNING +SVM)

	Healthy no-epileptic (Normal) To Interictal	healthy no-epileptic (Normal) To Ictal
ACC	94,50%	100%
SEN	93.25%	100%
SPE	95.75%	100%

The comparison with the previous studies reported in the literature is depicted in Table 5. From this table, the classification results obtained by the experiment of the proposed method (Y-W+Burg with Deep learning and SVM), they have achieved an excellent performance of (100.00% 100,00% 100.00%) accuracy, sensitivity, specificity, respectively in the sets (A and E). It can be concluded that the introduced method is better than the previous methods.

TABLE V. COMPARISON BETWEEN ACCURACIES OF THE PROPOSED METHOD AND OTHER METHODS FROM THE LITERATURE USING DIFFERENT SETS OF THE EPILEPTIC DATA.

Authors	Dataset	Methods	Accuracy
This work	A and E	AR Y-W+Burg +SVM	98.50%
		AR Y-W+Burg+deep learning +SVM	100%
	A and D	AR-Y-W+Burg +SVM	95.00%
		AR-Y-W+ Burg+deep learning +SVM	93.00%
Guo et al. [6]	A and E	ANN	99.66%
Subasi and [9]	A and E	SVM+PCA SVM+ICA SVM+LDA	98.75% 99.50% 100%
Lin, et al[10]	A and E B and E A and D	SSAE networks - softmax classifier	100% 100% 93.33%
Chandaka et al. [5]	Aand E	Cross-correlation aided SVM classifier signal	95.96%
Nicolaou, et al. [8]	A and E	Permutation Entropy (PE) and SVM	93.55%
	A and D		88.83%

IV. CONCLUSION

The current paper introduced a new framework based on AR model and deep learning to better classify epileptic seizures

from EEG signals. Compared with the previous studies that have been reported in the literature, this method is capable to provide more knowledge and can better distinguish between normal EEG signals and epileptic seizure. Also, the improvement of AR coefficients by deep learning is a powerful and effective technique that achieved a maximum classification accuracy of 100.00%. In addition, both experiments' results proved the ability of the proposed method to achieve a better average of classifications accuracies between 93.50 % and 100% depending on the use of a couple of data. Hence, it is recommended to employ this method in automatic detection of neurological disorders. The future work aims to apply the introduced method to other EEG datasets provided from subjects doing some tasks in order to recognize different mental tasks.

REFERENCES

[1] A. Alkan, E. Koklukaya, and A. Subasi. Automatic seizure detection in EEG using logistic regression and artificial neural network. Journal of Neuroscience Methods, 148(2), 167-176, 2005.

[2] H. Adeli, Z. Zhou and N. Dadmehr. Analysis of EEG records in an epileptic patient using wavelet transform. Journal of neuroscience methods, 123(1), 69-87, 2003.

[3] ATTIA, Abdelouahab, MOUSSAOUI, Abdelouahab, CHAHIR, Youssef. An EEG-fMRI Fusion Analysis Based on Symmetric Techniques Using Dempster Shafer Theory. Journal of Medical Imaging and Health Informatics, vol. 7, no 7, p. 1493-1501. 2017.

[4] N. Yalcin, G. Tezel, and C. Karakuzu. Epilepsy diagnosis using artificial neural network learned by PSO. Turkish Journal of Electrical Engineering & Computer Sciences, 23(2), 421-432, 2015.

[5] S. Chandaka, A. Chatterjee and S. Munshi. Cross-correlation aided support vector machine classifier for classification of EEG signals. Expert Systems with Applications, 36(2), 1329-1336,2009.

[6] L. Guo, D.J. Rivero, J.R. Rabunal and A. Pazos. Automatic epileptic seizure detection in EEGS based on line length feature and artificial neural networks. J. Neurosci. Meth. 191(1), 101–109 , 2010.

[7] L. Guo, D. Rivero, J. Dorado, C.R Munteanu and A. Pazos. Automatic feature extraction using genetic programming: an application to epileptic EEG classification. Expert Syst. Appl. 38(8), 10425–10436 , 2011.

[8] N. Nicolaou and J. Georgiou: Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Expert Syst. Appl. 39(1), 202– 209, 2012.

[9] A. Subasi, and M.I. Gursoy. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Systems with Applications, 37(12), 8659-8666, 2010.

[10] Q. Lin, S.Q. Ye, X.M. Huang., S.Y. Li, M.Z. Zhang, Y. Xue, and W.S. Chen. Classification of Epileptic EEG Signals with Stacked Sparse Autoencoder Based on Deep Learning. In International Conference on Intelligent Computing, 802-810. 2016

[11] S.G. Fabri, K.P. Camilleri, and T. Cassar. Parametric modeling of EEG data for the identification of mental tasks. INTECH Open Access Publisher, 2011.

[12] F. Shiman, S.H. Safavi, F.M. Vaneghi, M. Oladazimi and M.J. Safari, and F. Ibrahim .EEG feature extraction using parametric and non-parametric models. In Proceedings IEEE-EMBS International Conference on Biomedical and Health Informatics.2012

[13] K. Padmavathi, and K. Ramakrishna. Classification of ECG signal during Atrial Fibrillation using Burg's method. International Journal of Electrical and Computer Engineering, 5(1), 64. 2015

[14] A. Subasi, A. Alkan, E. Koklukaya, and M.K. Kiymik. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. Neural Networks, 18(7), 985-997.2005

[15] F. Shiman, S.H. Safavi, F.M. Vaneghi, M. Oladazimi, M.J. Safari and F. Ibrahim. EEG feature extraction using parametric and non-parametric models. In Proceedings of 2012 IEEE-EMBS International Conference on Biomedical and Health Informatics.

[16] A. Ouelli, B. Elhadadi, H. Aissaoui, and B. Bouikhalene. Epilepsy Seizure Detection Using Autoregressive Modelling and Multiple Layer Perceptron Neural Network. American Journal of Computer Science and Engineering, 2(4), 26.2015.

[17] J.P. Burg. A new analysis technique for time series data,"NATO Adv. Study Ins. on Signal Proc. with Emphasis on Underwater Acoustics, Netherlands, August 1968.

[18] D. J Krusiensk, D.J. McFarland and J.R Wolpaw, J. R. An evaluation of autoregressive spectral estimation model order for brain-computer interface applications. In Engineering in Medicine and Biology Society,2006.

[19] N. Kamel, A. Samraj, and A. Mousavi. Whitening of background brain activity via parametric modeling. Discrete Dynamics in Nature and Society, 2007.

[20] K. Vedavathi, K.S. Rao, and K.N. Devi . Unsupervised learning algorithm for time series using bivariate AR (1) model. Expert Systems with Applications, 41(7), 3402-3408.2014.

[21] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David and C.E. Elger. Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. Phys. Rev. E 64:061907. 2001

ITERATIVE COLLOCATION METHOD FOR VOLTERRA INTEGRO-DIFFERENTIAL EQUATIONS

Abstract. This paper deals with the numerical solution of nonlinear Volterra integro-differential equations by spline collocation. We applied the iterative collocation method to obtain an approximate solution without needed to solve any algebraic system. The analysis of the error has been discussed. Numerical examples are also presented to validate the theoretical analysis.

key words Volterra integro-differential equations, Collocation method, Iterative Method, Lagrange polynomials.

2010 AMS Subject Classification 45L05, 65R20.

1. Introduction

In this paper, we investigate an iterative collocation method for the following Volterra integro-differential equation

$$x'(t) = f(t) + \int_0^t K(t, s, x(s))ds, x(t_0) = x_0, t \in I = [0, T], \tag{1.1}$$

where the functions f, K are sufficiently smooth.

There are several numerical methods for approximating the solution of equation (1.1). For example, spectral methods, implicit RungeKutta methods, Galerkin methods, collocation methods, and Legendre wavelets series, (cf, e.g. [24, 26, 21, 2, 3, 4], and references therein).

The purpose of this paper is to solve equation (1.1) by the iterative collocation method.

This paper is concerned with the iterative collocation method to obtain an approximate solution for (1.1), our method presents some advantages:

- It provides a global approximation of the solution
- Without needed to solve any algebraic system

- High order of convergence
- Provides an explicit numerical solution and easy to be implemented.

The outlines of this paper is as follows. In section 2, the spline polynomial has been used to approximate equation (1.1) based on the iterative collocation method, error analysis has been discussed in section 3, section 4 is devoted to present some numerical examples, in the last section, we give a conclusion.

2. Description of the collocation method

We define the real polynomial spline space of degree $m+1$ as follows:

$$S_{m+1}^{(1)}(\Pi_N) = \{u \in C^1(\mathrm{I}, \mathbb{R}) : u_n = u/\sigma_n \in \pi_{m+1}, n = 0, ..., N-1\}.$$

It holds for any $u \in S_{m+1}^1(I, \Pi_N)$ that

$$u_n'(t_n + sh) = L_0(v)u_{n-1}'(t_n) + \sum_{j=1}^{m} L_j(v)u_n'(t_{n,j}) \tag{2.1}$$

Now, we approximate x by $u \in S_{m+1}^1(I, \Pi_N)$ such that $u'(t_{n,j})$ satisfy the following nonlinear system,

$$\begin{aligned} u_n'(t_{n,j}) = f(t_{n,j}) &+ h\sum_{p=0}^{n-1} K(t_{n,j}, t_p, u_p(t_p)) + h^2 \sum_{p=0}^{n-1} b_0 K'(t_{n,j}, t_p, u_p(t_p)) \\ &+ h^2 \sum_{p=0}^{n-1}\sum_{v=1}^{m} b_v K'(t_{n,j}, t_{p,v}, u_p(t_{p,v})) + hc_j K(t_{n,j}, t_n, u_{n-1}(t_n)) \\ &+ h^2 a_{j,0} K'(t_{n,j}, t_n, u_{n-1}(t_n)) + h^2 \sum_{v=1}^{m} a_{j,v} K'(t_{n,j}, t_{n,v}, u_n(t_{n,v})), \end{aligned} \tag{2.2}$$

for $n = 0, ..., N-1$, $j = 1, ..., m$ where $u_{-1}'(t_0) = x'(0) = f(0)$ and $u_{-1}(t_0) = x(0)$. Since the above system is nonlinear, we will use an iterative collocation solution $u^q \in S_{m+1}^1(I, \Pi_N), q \in \mathbb{N}$, to approximate the solution of (1.1) such that

$$(u_n^q)'(t_n + sh) = L_0(s)(u_{n-1}^q)'(t_n) + \sum_{j=1}^{m} L_j(s)(u_n^q)'(t_{n,j}), s \in [0, 1], \tag{2.3}$$

and

$$u_n^q(t_n+sh) = u_{n-1}^q(t_n)+hB_0(s)(u_{n-1}^q)'(t_n)+h\sum_{j=1}^{m} B_j(s)(u_n^q)'(t_{n,j}), s \in [0,1], \quad (2.4)$$

where the coefficients $(u_n^q)'(t_{n,j})$ are given by the following formula:

$$\begin{aligned}(u_n^q)'(t_{n,j}) =& f(t_{n,j}) + h\sum_{p=0}^{n-1} K(t_{n,j}, t_p, u_p^q(t_p)) + h^2\sum_{p=0}^{n-1} b_0 K'(t_{n,j}, t_p, u_p^q(t_p)) \\ &+ h^2\sum_{p=0}^{n-1}\sum_{v=1}^{m} b_v K'(t_{n,j}, t_{p,v}, u_p^q(t_{p,v})) + hc_j K(t_{n,j}, t_n, u_{n-1}^q(t_n)) \\ &+ h^2 a_{j,0} K'(t_{n,j}, t_n, u_{n-1}^q(t_n)) + h^2\sum_{v=1}^{m} a_{j,v} K'(t_{n,j}, t_{n,v}, u_n^{q-1}(t_{n,v})),\end{aligned} \quad (2.5)$$

such that $(u_{-1}^q)'(t_0) = f(0)$ and $u_{-1}^q(t_0) = x_0$ for all $q \in \mathbb{N}$ and the initial values $(u_n^0)'(t_{n,j}) \in J$ (J is a bounded interval).

The above formula is explicit and the approximate solution u^q is given without needed to solve any algebraic system.

In the next section, we will prove the convergence of the approximate solution u^q to the exact solution x of (1.1), moreover, the order of convergence is m for all $q \geq m$.

3. Convergence analysis

In our convergence analysis, we study the following linear Volterra integro-differential equation

$$x'(t) = f(t) + \int_0^t K(t,s)x(s)ds, t \in I = [0,T], \quad (3.1)$$

The following three lemmas will be used in this section.

Lemma 3.1. [20] *Assume that $(\alpha_n)_{n\geq 1}$ and $(q_n)_{n\geq 1}$ are given non-negative sequences and the sequence $(\varepsilon_n)_{n\geq 1}$ satisfies*

$\varepsilon_1 \leq \beta$ and

$$\varepsilon_n \leq \beta + \sum_{j=1}^{n-1} q_j + \sum_{j=1}^{n-1} \alpha_j \varepsilon_j, \ n \geq 2.$$

Then

$$\varepsilon_n \leq \left(\beta + \sum_{j=1}^{n-1} q_j\right) \exp\left(\sum_{j=1}^{n-1} \alpha_j\right) \ , \ n \geq 2.$$

Lemma 3.2. [1] *Assume that the sequence* $\{\varepsilon_n\}_{n\geq 0}$ *of nonnegative numbers satisfies*

$$\varepsilon_n \leq A\varepsilon_{n-1} + B\sum_{i=0}^{n-1} \varepsilon_i + K, \quad n \geq 0,$$

where $A,\ B$ *and* K *are nonnegative constants, then*

$$\varepsilon_n \leq \frac{\varepsilon_0}{R_2 - R_1}\left[(R_2 - 1)R_2^n + (1 - R_1)R_1^n\right] + \frac{K}{R_2 - R_1}\left[R_2^n - R_1^n\right],$$

where

$$\begin{aligned} R_1 = &\quad \left(1 + A + B - \sqrt{(1-A)^2 + B^2 + 2AB + 2B}\right)/2, \\ R_2 = &\quad \left(1 + A + B + \sqrt{(1-A)^2 + B^2 + 2AB + 2B}\right)/2, \end{aligned} \tag{3.2}$$

therefore, $0 \leq R_1 \leq 1 \leq R_2$.

The following result gives the existence and the uniqueness of a solution for the linear system (3.3).

Lemma 3.3. *For sufficiently small* h*, the linear system* (3.3) *defines a unique solution* $u \in S^1_{m+1}(I, \Pi_N)$ *which is given by* (2.5).

The following result gives the convergence of the approximate solution u to the exact solution x.

Theorem 3.4. *Let* f, K *be* $m+2$ *times continuously differentiable on their respective domains. If* $-1 < R(\infty) = (-1)^m \prod_{l=1}^{m} \frac{1-c_l}{c_l} < 1$*, then, for sufficiently small* h*, the collocation solution* u *converges to the exact solution* x*, and the resulting errors functions* $e^{(v)} := x^{(v)} - u^{(v)}$ *for* $v = 0, 1$ *satisfies:*

$$\|e^{(v)}\|_{L^\infty(I)} \leq Ch^{m+1},$$

for $v = 0, 1$ *and* C *is a finite constant independent of* h.

The following result gives the convergence of the iterative solution u^q to the exact solution x.

Theorem 3.5. *Consider the iterative collocation solution $u^q, q \geq 1$ defined by (2.8), if $-1 < R(\infty) = (-1)^m \prod_{l=1}^{m} \frac{1-c_l}{c_l} < 1$, then for any initial condition $(u')^0(t_{n,j}) \in J$, the iterative collocation solution $u^q, q \geq 1$ converges to the exact solution x for sufficiently small h. Moreover, the following errors estimates hold*

$$\|u^q - x\| \leq \gamma_1 \beta^q h^{2q} + \gamma_2 \beta^q h^{m+1+2q} + \beta_2 h^{m+1}$$

and

$$\|(u^q)' - x'\| \leq \beta_1 h^{2q} + \beta_2 h^{m+1+2q} + \beta_2 h^{m+1}.$$

where $\beta, \beta_1, \beta_2, \gamma_1, \gamma_2$ are finite constants independent of h and q.

4. Numerical Examples

Example 4.1. *([7, 8, 25]) Consider the following nonlinear Volterra integral equation*

$$x'(t) = 2\sin(t)\cos(t) + 3\int_0^t \cos(t-s)(x(s))^2 ds, \ t \in [0,1],$$

where $x(x) = \cos(x)$ is the exact solution.
The absolute errors for $N = m = q = 4$ at $t = 0, 0.1, ..., 1$ are displayed in Table 1. We used the collocation parameters $c_i = \frac{i}{m+1} + \frac{1}{4}, i = 1, ..., m$ and $R(\infty) = -\frac{11}{1989}$. The numerical results of the present method are considerable accurate in comparison with the numerical results obtained by [7, 8, 25].

5. Conclusion

This paper has considered the iterative collocation method approximation approach for solving nonlinear Volterra integral equations. The method is easy to implement and has high order of convergence. The convergence of the presented algorithm is proved and an error estimate is established. Iterative collocation method can be extended to higher order integro- differential equations. Thus a possible area of future research is the application of the iterative collocation method method to

TABLE 1. Comparison of the absolute errors of Example 4.1

t	*Method in* [25] $N = 16$	*Method in* [8] $N = 32$	*Method in* [7] $N = 32$	*Present method* $N = 4$
0.0	0.0	0.0	$----$	0.0
0.1	4.43×10^{-4}	4.49×10^{-4}	1.09×10^{-3}	9.27×10^{-10}
0.2	2.22×10^{-4}	2.42×10^{-4}	7.25×10^{-4}	4.19×10^{-8}
0.3	1.22×10^{-4}	1.62×10^{-4}	8.42×10^{-4}	1.10×10^{-7}
0.4	1.34×10^{-4}	2.00×10^{-4}	3.56×10^{-3}	1.97×10^{-7}
0.5	4.29×10^{-4}	3.38×10^{-4}	7.59×10^{-3}	3.10×10^{-7}
0.6	1.77×10^{-4}	6.10×10^{-5}	5.29×10^{-3}	4.20×10^{-7}
0.7	4.54×10^{-4}	3.22×10^{-4}	1.94×10^{-3}	5.33×10^{-7}
0.8	5.75×10^{-4}	4.35×10^{-4}	2.34×10^{-3}	6.52×10^{-7}
0.9	5.82×10^{-4}	4.47×10^{-4}	1.69×10^{-4}	7.29×10^{-7}
1.0	9.15×10^{-4}	8.00×10^{-4}	$----$	7.84×10^{-7}

higher dimensional problems. One could also investigate the application of iterative collocation method to singular PDE's.

REFERENCES

[1] E. Hairer, C. Lubich, S. P. Nϕrsett, Order of convergence of one-step methods for Volterra integral equations of the second kind, SIAM J. Numer. Anal 20 (1983), 569-579.

[2] K.E. Atkinson, The Numerical Solution of Integral Equations of the Second Kind. Cambridge University Press, Cambridge, 1997.

[3] R. Kress, Linear Integral Equations. Springer-Verlag, NewYork, 1999.

[4] P.K. Kythe and P. Puri, Computational methods for linear integral equations. Birkhauser-Verlag, Springer, Boston, 2002.

[5] L. Hacia, Iterative-Collocation Method for Integral Equations of Heat Conduction Problems, Numerical Methods and Applications, (2006), 378-385.

[6] H. Brunner, iterated collocation methods for volterra integral equations with delay arguments, Mathematics of Computation,62 (1994), 581-599.

[7] E. Babolian, Z. Masouri, and S. Hatamzadeh-Varmazyar, New direct method to solve nonlinear Volterra-Fredholm integral and integro-differential equations using operational matrix with block-pulse functions, Prog. in Electromag. Research 8 (2008), 59-76.

[8] E. Babolian, Z. Masouri, and S. Hatamzadeh-Varmazyar, Numerical solution of nonlinear Volterra-Fredholm integro-differential equations via direct method using triangular functions, Comp. Math. Appl. 58 (2009), 239-247.

[9] N. Bildik, A. Konuralp, S. Yalçnbas, Comparison of Legendre polynomial approximation and variational iteration method for the solutions of general linear Fredholm integro-differential equations, Comput. Math. Appl. 59 (2010) 19091917.

[10] S.H. Wang, J.H. He, Variational iteration method for solving integro-differential equations, Phys. Lett. A 367 (2007) 188191.

[11] J.I. Ramos, Iterative and non-iterative methods for non-linear Volterra integro-differential equations, Appl. Math. Comput. 214 (2009) 287296.

[12] R. A. Frazer, W. P. Jones, and S. W. Skan, Approximations to Functions and to the Solutions of Differential Equations. Gt. Brit. Aero. Res. Council Reut and Memo: renrinted in Gt. Brit. Air Ministry Aero. Res. Comm. Tech. Report Vol. 1 (1937), 517-549.

[13] M. Costabel and J. Saranen, Spline Collocation for Convolutional Parabolic Boundary Integral Equations, ACM Numer. Math. 84 (2000), 417-449.

[14] W.H. Huang and R.D. Russell, A Moving Collocation Method for Solving Time Dependent Partial Differential Equation, SIAM J.Appl. Numer. Math. 20 (1996), 101-116.

[15] P. Sridhar, Implementation of the One Point Collocation Method to an Affinity Packed Bed Model, Indian Chem. Eng. Sec. 41(1) (1999), 39-46.

[16] M. Thamban Naira and V. Sergei Pereverzevb, Regularized Collocation Method for Fredholm Integral Equations of the First Kind, Journal of Complexity. 23 (2007), 454-467.

[17] J.P. Coleman, S.C. Duxbury, Mixed collocation methods for $y = f(x, y)$, J. Comput. Appl. Math. 126 (2000), 47-75.

[18] H. Brunner, A. Makroglou, R.K. Miller, Mixed interpolation collocation methods for first and second order Volterra integro-differential equations with periodic solution. Appl. Numer. Math. 23 (1997), 381-402.

[19] H. Brunner, Collocation Methods for Volterra Integral and Related Functional Differential Equations. Cambridge University Press, Cambridge (2004).

[20] H. Brunner and P. J. van der Houwen, The numerical solution of Volterra equations, CWI Monogr., vol. 3, North-Holland, Amsterdam, 1986.

[21] H. Brunner, Implicit RungeKutta methods of optimal order for volterra integro-differential equations. Math. Comput.42(165) (1984), 95109 .

[22] H. Liang, H. Brunner, On the convergence of collocation solutions in continuous piecewise polynomial spaces for Volterra integral equations. BIT Numerical Mathematics. 56 (2016), 1339-1367.

[23] H. Laib, A. Bellour and M. Bousselsal, Numerical solution of high-order linear Volterra integro-differential equations by using Taylor collocation method. International Journal of Computer Mathematics, 96(5) (2019), 1066-1085.

[24] Y.-J. Jiang, On spectral methods for Volterra-type integro-differential equations. J. Comput. Appl. Math.230(2) (2009), 333340.

[25] S. Ul-Islam, I. Aziz, M. Fayyaz, A new approach for numerical solution of integrodifferential equations via Haar wavelets. International Journal of Computer Mathematics. 90(3) (2013), 1971-1989.

[26] T. Tang, X. Xiang, J. Cheng, On spectral methods for Volterra integral equations and the convergence analysis. J. Comput. Math.26(6) (2008), 825837.

[27] S. Yalçinbaşa, M. Sezer, The approximate solution of high-order linear Volterra-Fredholm integro-differential equations in terms of Taylor polynomials, Appl. Math. Comput. 112 (2000), 291-308.

[28] Ş. Yüzbaşi, N. Şahin, M. Sezer, Bessel polynomial solutions of high-order linear Volterra integro-differential equations, Computers and Mathematics with Applications, 62 (2011), 1940-1956.

Ontology-Based Context Modeling for Adaptive Learning

Ouissem Benmesbah
RSI Laboratory
Badji Mokhtar University
Annaba - Algeria
ouissa2007@yahoo.fr

Lamia Mahnane
RSI Laboratory
Badji Mokhtar University
Annaba - Algeria
mahnane_lamia@yahoo.fr

Mohamed Hafidi
RSI Laboratory
Badji Mokhtar University
Annaba - Algeria
mhhafidi@yahoo.fr

Abstract— **Adaptive learning is taking place in extremely various and rich environments; Context-awareness is a fundamental ingredient of this kind of learning, consequently, the ability to model and to recognize the user's context is necessary to effectively adapt pedagogical content and learning activities to the learner's needs, goals and environmental characteristics. The requirement to define context more precisely and in a uniform way has been identified by several researchers. A general and precise definition of context can facilitate the identification of what does and does not constitute context and can enable reuse and share of contextual data over applications. To this extent this work aims to build reference ontology for adaptive and personalized mobile learning environment called RefOnto.**

Keywords— Context model, reference model, Ontology, Context-aware Learning, Hierarchical model, adaptive learning

I. INTRODUCTION

Mobile, smart devices supporting emerging pervasive applications will constitute a significant part of future technologies by providing highly proactive services requiring continuous monitoring of user related contexts. In technology-enhanced learning (TEL), the support for context-awareness is essential so that it can make learning contextual and improve learning experiences, allowing learners to carry out daily activities anytime and anywhere [1][2].

The need to define and model context more precisely and in a consistent way has been identified by several researchers. A precise definition of context can facilitate the identification of what does and does not constitute context and can enable reuse and exchange of contextual data across applications [20].

The challenge in context modeling is to identify those pieces of contextual information that are considered most relevant to the majority of existing adaptive applications; these applications have been developed specifically for one selected scenario. What yet is missing is a reference model that enables the methodical development of adaptive mobile applications that support a wide variety of educational scenarios [4][1][5][3].

The research contribution of this paper is threefold.

1. First, we present a context framework which identifies relevant context dimensions for TEL applications. The proposed model is developed by consolidating the various context dimensions used in the existing context aware Learning applications and organizing them into an appropriate structure.

2. Second, Implement the proposed framework using Ontology in owl format.

3. Finally, Evaluate the proposed ontology using a set of scenarios and metrics to assess the applicability and the richness of the model.

The remainder of this paper is articulated as follows, Section 2 presents a theoretical framework to introduce the basic concepts in the field of context awareness. In Section 3, we analyze several existing context models in Context aware Learning field, in order to conclude which characteristics should be involved in RefOnto, then the design of the proposed model is provided in Section 4. Section 5, demonstrates the practicality and the richness of our ontology through several scenarios and a comparative study is conducted with a golden standard work. The conclusion of this part of the research is introduced in Section 6.

II. THEORETICAL FRAMEWORK

This section focuses on the study of the basic concepts in the field of context awareness:

A. Context and Context awareness

Dey expresses that for the computation [7]: ''Context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application". Dey [7] also defined a system as context-aware if "it uses context to provide relevant information and/or services to the user, where relevancy depends on the user's task".

B. Context awareness and content adaptation

The strengths of virtual learning systems, targeted at mobile devices, can be improved by mingling context awareness with content adaptation. The context awareness is formed by data regarding users and their environment, such as time, location, device characteristics, learning objective, and preferences, among others. Content adaptation, accordingly, can personalize the learning object to meet this context [10].

Ontology is considered a core component of semantic web, which supports the process of personalization and improves the capabilities of adaptive learning systems [2][8]. Based on Ontologies, a system could possess the ability to provide semantic context description, context reasoning, knowledge sharing, context classifications, context dependency [3][6]. In the scope of adaptive and smart systems, there is a trend on using ontology to promote adaptive services directed to

education [10]. In the Related works section, we will focus only on ontology-based works; this choice is justified by the popularity of ontology and its leading role in personalization and adaptive learning environment [9][34].

III. RELATED WORKS

This section provides an overview of the context dimensions used within the field of Context modeling for Context-aware and adaptive Learning.

TABLE I. OVERVIEW OF ONTOLOGY BASED U-LEARNING APPLICATIONS

Context Dimension	14	18	15	19	5	21	24	2	23	22
Profile	*	*	*	*	*	*	*	*	*	
Location	*	*	*		*	*	*			*
Time	*	*	*			*	*			*
Needs		*							*	
Mobility					*					
Physical Cdt			*			*	*			
Technology	*			*		*	*			*
Concentration Level										
Preferences		*						*		
Knowledge Level		*					*	*		
Learning Style		*						*		
Satisfaction									*	
Activity Type								*		
Teaching method								*		

This study allowed us to identify the most commonly used context dimensions within the field of context aware and adaptive learning, but also to reveal the following gaps:

- Existing works integrate a part of the context, what confirm that there is no a reference model that represent learning context in its entirety.
- Some works use only context related to the learner physical sate (Location, Time, technology, and mobility) to the detriment of the context that reflects the learning profile (Knowledge level, learning style, learner feedback, …). Other works combine the two types of information, but they integrate few dimensions related to the learning side.
- We notice From Table 1 that some contextual information are common to several works, which mean that they appear frequently within various scenarios. This finding has not been exploited to promote the reuse of the existing context model which can significantly reduce the time spent on design.

The previous gaps led us to confirm the objective of this work, which consists in designing a reference model for learners' context called RefOnto, which aims at collecting the most used information in the field of adaptive and context aware learning, offering a unified vision that takes into account on one hand the particular situation of the learner in a physical environment (location, time, physical state, technology, ...) which is general information common to all context aware applications and on the other hand the information that characterizes its educational profile (knowledge level, learning style, ...).

IV. PROPOSED REFERENCE CONTEXT MODEL FOR CONTEXT-AWARE LEARNING

A. *PROPOSED FRAMEWORK*

Ontology is considered a core component of semantic web, which supports the learning process personalization and improves the competencies of learning systems [2]. Based on ontology, a system could possess the abilities like semantic context description, context reasoning, knowledge sharing, context classifications, context dependency, etc [3][6].

In order to create a reference context model adapted to any learning scenario, it is necessary to develop an ontology that captures generic concepts in a Top level and be able to extend the specific information of a context in a hierarchical manner [6].

In this light, our context ontology is characterized by two hierarchical levels (see Fig. 1). The first one is general and common domain ontology, that will be shared among different particular domains and the second one is particular domain ontology which, inherits and specifies the common ontology and captures contexts of a specific domain (transportation, medical service, etc.), in our case the Learning domain.

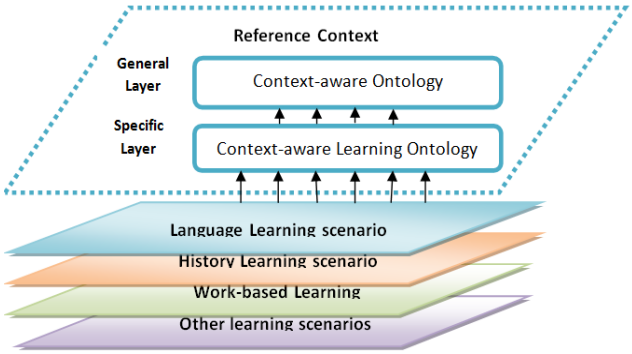


Fig. 1 A Reference context model framework

B. *Methodology for the development of the proposed ontology*

We are basing our development process on Noy and McGuiness's method [17].

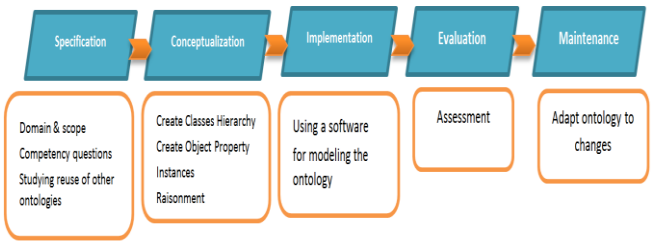


Fig. 2. Methodology used for developing RefOnto

C. Description of the ontology used to build the proposed model

The proposed ontology RefOnto is created from scratch. We determined its hierarchy, set of classes, subclasses and their relationships. Protégé is used to build our ontology. It is considered the most popular, widely used, freely available platform and managing knowledge bases, ontologies and terminologies in a wide range of application domain [6].

A general representation of the proposed context model is illustrated in Fig. 3.

Four steps are followed to create our ontology, step1 creates classes and class hierarchy, step2 determines object properties, step3 determines ontology instances, and step 4 is dedicated to ontology reasoning. In this paper, only the first two steps are realized.

Step1. Creates a class hierarchy

Classes and properties that form the proposed reference model are depicted from the Section 3; therefore, these entities are the most common elements of the major works in the field of context-aware learning. All concepts related to learners' context are determined and organized in two layers:

a. General layer: Defines five main classes: User, Task, Environment, Technology and Resources.

b. Specific layer: Extends four main classes from the higher level as shown in Fig. 3:

- **Learner state:** Includes information about Learner Identity, Mobility, Feedback, preferences, learning objective, learning style, knowledge level. Relevant standards and specifications for partial representations of learner data are IMS LIP, IMS ePortfolio, IMS Enterprise, IEEE RCD, FOAF, and HR-XML [20].
- **Learning Task**: Reflects the tasks, objectives or actions of the user, it has the following subclasses: Identification, Difficulty level, Task duration, learning concept.
- **Learning environment:** Describes information that characterize the learners' environment, it includes: Time, Location and Physical conditions.
- **Learning technology:** Device and Network classes. The prevalent standards for describing Technology context are: are W3C Composite Capabilities/Preferences Profile (CC/PP), User Agent Profile (UAprofile), developed by OMA, and the Usage Environment Description (UED) standard which was standardized within MPEG-21 Digital Item Adaptation [20].
- **Learning Resources:** Captures relevant characteristics of physical or virtual resources. LOM (Learning Object Metadata), Dublin Core and MPEG-6 are standards for the description of the learning resources.

Step2. Determine object properties,

More details about relationships among classes are presented in Table 2.

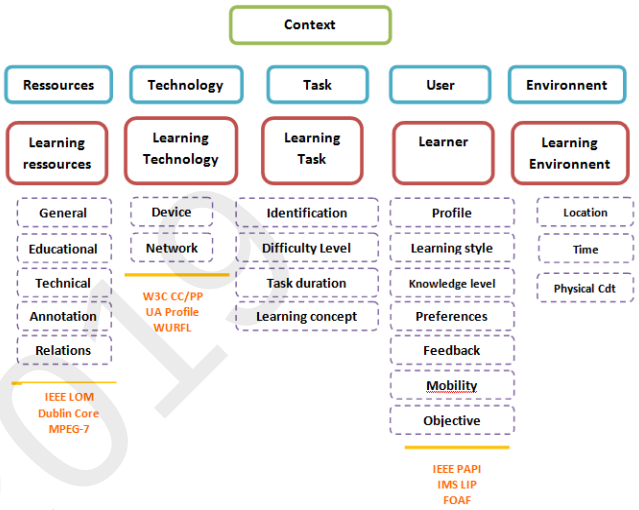


Fig. 3. Hierarchical design of the proposed generic context model

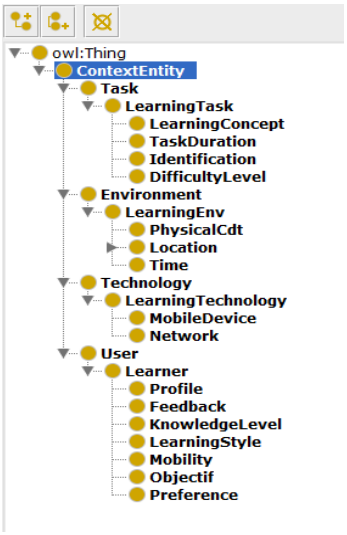


Fig. 4. Hierarchical design implemented in Protégé

TABLE II. OBJECT PROPERTIES IN THE PROPOSED GENERIC CONTEXT MODEL

Object property	Domain	Range	Meaning
ActiveIn	User	Environment	User is active in an Environment
Uses	User	Technology	User uses a technology
EngagedIn	User	Task	User is engaged in an activity
hasConnection	MobileDevice	Network	A mobile device has a network connection
isRelatedTo	Resource	Task	A resource is related to a task
UsedIn	Technology	Environment	A technology is used in an environment

V. EVALUATION AND RESULTS:

Ontology evaluation is an important issue that must be addressed if ontologies are to be widely adopted in the semantic web. Users facing a multitude of ontologies need to have a way of assessing them and deciding which one best fits their requirements the best [25]. To accomplish this task, we employed two approaches:

A. Senarios

In this section, we propose a set of scenarios in order to evaluate the behavior of the proposed ontology. We chose to present just three scenarios among several possible scenarios.
Semantic web Rule Language is an important formalism for expressing knowledge in the form of rules. SWRL is used to define inference rules in knowledge models represented in OWL in a semantically consistent way [9]. In this work, our Ontology based model has been enriched with a set of rules, to be used in different ways, for example to enable inference mechanisms to automatically infer the learning style of learners. This property plays a major role to create a successful and effective adaptation. Learning style refers to how a learner uses his/her senses to learn. The learning style category can be determined by answering the ILS questionnaire. For example, if ILS questionnaire answers specify that the learner belongs to "Active" category (See Fig 5), context model ontology should be updated using the following rule:

Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Active') → hasLearningStyle (?x, 'Active')
Similar sules are applied for others results of the ILS Questionnaire.

Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Reflective') → hasLearningStyle (?x, 'Reflective')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Intuitive) → hasLearningStyle (?x, 'Intuitive')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Sensing') → hasLearningStyle (?x, 'sensing')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Verbal') → hasLearningStyle (?x, 'Verbal')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Visual') → hasLearningStyle (?x, 'Visual')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Sequencial') → hasLearningStyle (?x, 'Sequencial')
Learner (?x) ^ hasLearningStyle (?x,?y) ^ is (?Y, 'Global'') → hasLearningStyle (?x, 'Global')

Fig. 5. Rule 1

Another use of SWRL Rules is to infer the Difficulty level of the learning activity based on the Knowledge level of the learner (Fig 6):

Learner (?x) ^ hasKnowledgeLevel (?x, ''Beginner'') ^ LearningActivity (?y) ^ EngagedIn (?x, ?y) ^ → hasDifficultyLevel (?y, ''Low'')

Learner (?x) ^ hasKnowledgeLevel (?x, ''Medium'') ^ LearningActivity (?y) ^ EngagedIn (?x, ?y) ^ → hasDifficultyLevel (?y, ''Medium'')

Learner (?x) ^ hasKnowledgeLevel (?x, ''Advanced'') ^ LearningActivity (?y) ^ EngagedIn (?x, ?y) ^ → hasDifficultyLevel (?y, ''high'')

Fig. 6. Rule 2

The last example of using SWRL Rules in our proposed model aims to infer the type of learning activity based on the learning style of the learner. According to Table5, there is a relation between learners learning style and the proposed learning Activity type. Taking as an example, Sequential learner, who likes to learn in specific, sequenced and small Steps and applies logical stepwise paths in solving problems, the most suitable learning activity type for this category of learners, is case Study. Whereas, if the learner has visual learning style, he/she prefer information, which is presented visually. In this case the suitable learning activities are: visualization exercises, problem solving and concept map (See Fig7).

Learner (?x) ^ hasLearningStyle (?x, 'Sequential') ^ LearningActivity (?y) ^ EngagedIn (?x, ?y) → hasActivityType (?y, 'Case Study')

Learner (?x) ^ hasLearningStyle (?x, 'Visual) ^ LearningActivity (?y) ^ EngagedIn (?x, ?y) → hasActivityType (?y, 'concept map')

Fig. 7. Rule3

B. Metric-based evaluation

Metrics

We used the methodology described in [10], which is based on FOEval model [11] to evaluate the proposed context ontology. FOEval [11] consists in a group of metrics that assist the evaluation of local or remote ontologies. We employed the metrics richness as proposed in FOEval. Ontology richness can be measured on different levels:

- Relation richness (rR)
- Attribute richness (aR)
- Ontology richness (oR)
- Subclass richness (sR)

Relation richness (rR): This metric reflects the diversity of relations and placement of relations in the ontology.
The ontology that has more relationships (composition), instead of inheritances (specializations) is considered richer than the taxonomy with the opposite characteristic.

Attribute richness (aR): Denotes the amount of information stored by an ontology. The more attributes are defined in the Ontology, the better will be the knowledge that the ontology represents.

Ontology richness (oR): This metric can be used in comparison with other ontologies, in order to determine how the value of oR is significant.

Subclass richness (sR): Is a good indication of how well knowledge is grouped into different categories and subcategories in the ontology.
The metrics values are obtained using a tool called OntoMetric [26], which is used to validate and display statistics about an ontology described in OWL.

Comparative analysis

As a golden standard, we used the ontology described in [10]. Our choice is based on three reasons: The first one is that, both works have many similarities in conceptual terms, both works propose an ontology based context model for adapting the educational content to the learners' context. The second reason is that Abech[10] has integrated in his work, an evaluation section by using the same metrics described in Foeval[11]. This will allow us to exploit directly the data of [10]. The work presented in [10] in among one of the few works that provide access to the developed OWL files. The metrics in Table 3 are extracted from the software Protégé.
We notice from Table 1 that RefOnto exceeds the ontology presented in [10] in all the metrics cited in the table (number of object property, sub object properties, data property, classes and sub classes) , this can be justified by the fact that our ontology defined as a reference model covers a wider field, integrating more concepts and relationships as well as attributes that appear frequently in several works.

TABLE III. METRICS OF REFONTO AND THE GOLDEN STANDARD WORK

	[10]	RefOnto
Total number of Object property	15	60
Total number of sub Object property	2	44
Total number of Data property	25	84
Total number of classes	28	66
Total number of sub classes	2	65

Table 4 summarizes the metric in concordance with the FOEval model [11] comparing the values of the proposed ontology and the ontology described in [10].

TABLE IV. METRICS OF REFONTO AND THE GOLDEN STANDARD WORK

	[10]	RefOnto
rR	0.88	0.51
aR	0.89	1.27
oR	1.77	1.78
sR	0.06	0.98

The results for RefOnto are 0.51 for rR and 1.27 points for aR. These two results indicate that the context ontology defined in this work is richer in attributes than relations. The value of aR obtained in our ontology is better if compared to the work presented in [10], as it's defined in Foeval[11], the more attributes are defined in ontology, the better will be the knowledge that the ontology represents.
The value of the metric rR is the smaller than [10], it means that the number of hierarchical relations defined in our ontology is more interesting than the number of non-hierarchical relations, this can be explained by the fact the proposed ontology is an hierarchical model based on 2 levels so the relationships that link the concepts of the two levels are hierarchical in nature.
Adding both metrics we can obtain the oR which is 1.78 points. oR value is determinant to affirm that the compared ontology is richer and has more information (contains more attributes and relationships). We can notice from Table 4, that the value of oR of RefOnto is more interesting than [10], this is because the scope of the reference ontology is wider than the scope of golden standards ontology.
The Subclass Richness (sR) indicates how well knowledge is grouped into different categories and subcategories in the ontology. According to table 4, the value of sR is 0.98; this result indicates that RefOnto, exceeds [10] in the distribution of knowledge between different concepts and sub concepts. Moreover, it indicates that the ontology represents detailed knowledge. Moreover, an ontology with a high sR would be of a horizontal nature, which means that the proposed ontology represents a wide range of general knowledge.

VI. CONCLUSION AND PERSPECTIVES

The use of ontology for context modeling in Context-aware learning environments is a growing research area. It has been pointed a major issue, All models of the reviewed literature deal only with a subset of context information that is of interest in a Context-aware learning setting. As such, there is no proposal of a reference context model for this area. For addressing this issue, this work has presented a proposition for reference context ontology to unify contextual dimensions in this area. Furthermore, we have evaluated the proposed ontology by showing several scenarios. These scenarios showed the possibility of representing different situations applying rules to the proposed ontology. Additionally, the metrics Richness are used to allow a comparison of the proposed ontology with a golden standard work.
This contribution is still in progress, it needs more enrichment and had to be evaluated by experts.
Extra research should contain the last two steps of creating the proposed ontology (ontology instances and ontology reasoning).

REFERENCES

[1] G. Gabriela, E. Durán, and A. Amandi, "Context ontologies in ubiquitous learning environments". Ibero-American Conference on Artificial Intelligence. Springer, Cham, 2016.

[2] S. Ouf, M. AbdEllatif, S. EzzatSalama, and Y. Helmy, "A proposed paradigm for smart learning environment based on semantic web." Computers in Human Behavior, 2017, pp. 796-818.

[3] Y. Chuantao, B. Zhang, B. David, and Z. Xiong, "A hierarchical ontology context model for work-based learning". Frontiers of Computer Science, vol. 9(3), pp. 466-473, 2015

[4] A. A. Economides, "Adaptive context-aware pervasive and ubiquitous learning". International Journal of Technology Enhanced Learning, vol. 1 (3), 2005, pp. 169-192.

[5] S. Ennouamani, and Z. Mahani, "Designing A Practical Learner Model For Adaptive And Context-Aware Mobile Learning Systems". IJCSNS 18, no. 4 (2018): 84.

[6] J. Aguilar, M. Jerez, and T. Rodríguez, "CAMeOnto: Context awareness meta ontology modeling", Applied computing and informatics, vol 14, no. 2, 2018, pp. 202-213.

[7] A. K. Dey, G. D. Abowd, and D. Salber, "A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications". Human–Computer Interaction, 2001, vol. 16, no. 2, pp. 97-166

[8] M. Al-Yahya, R. George, and A. Alfaries, "Ontologies in E-learning: review of the literature." International Journal of Software Engineering and Its Applications, 2015, vol 9, no. 2, pp. 67-84.

[9] J. Ye, D. S. Dasiopoulou, G. Stevenson, G. Meditskos, E. Kontopoulos, I. Kompatsiaris, and S. Dobson, "Semantic web technologies in pervasive computing: A survey and research roadmap", Pervasive and Mobile Computing, 2015, vol. 23, pp. 1-25.

[10] M. Abech, C. A. Costa, J. L. Barbosa, S. J. Rigo, and R. R. Righi, "A model for learning objects adaptation in light of mobile and context-aware computing." Personal and Ubiquitous Computing, 2016, vol. 20(2), pp. 167-184.

[11] A. B. Bouiadjra, and S. M. Benslimane, "FOEval: Full ontology evaluation." In Natural Language Processing andKnowledge Engineering (NLP-KE), 2011 7th International Conference, 2011, pp. 464-468. IEEE

[14] A. Harchay, L. Cheniti-Belcadhi, and R. Braham, "A Context-aware Approach for Personalized Mobile Self-Assessment." J. UCS, 2015, vol21(8), pp. 1061-1085.

[15] B. Curum, C. P. Gumbheer, K. K. Kavi, and R. Cunairun, "A content-adaptation system for personalized m-learning." In Next Generation Computing Applications (NextComp), 2017 1st International Conference on, pp. 121-128. IEEE.

[16] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "METHONTOLOGY: From Ontological Art Towards Ontological Engineering," in Proceedings of the Ontological Engineering AAAI-97 Spring Symposium Series, Stanford University, EEUU, 1997.

[17] N. F. Noy and D. L. Mcguinness, "Ontology Development 101: A Guide to Creating Your First Ontology," 2001.

[18] G. W. Musumba and R. D. Wario, "Towards a Context-Aware Adaptive e-Learning Architecture". In Annual Conference of the Southern African Computer Lecturers' Association, 2018, pp. 191-206,Springer, Cham

[19] B. Bouihi and M. Bahaj, "An ontology-based architecture for context recommendation system" in E-learning and mobile-learning applications. In 2017 International Conference on Electrical and Information Technologies (ICEIT), 2017, pp. 1-6, IEEE.

[20] K. Verbert, N. Manouselis, X. Ochoa, M. Wolpers, H. Drachsler, I. Bosnic and E. Duval, "Context-aware recommender systems for learning: a survey and future challenges.", *IEEE Transactions on Learning Technologies*, 2012, *5*(4), 318-335.

[21] S. Baccari and M. Neji, "Design for a context-aware and collaborative mobile learning system". In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), 2016, pp. 1-6, IEEE

[22] H. Supic, "A model of a case-based approach to context-aware content sequencing" in Mobile learning environments. In 2016 International Conference on Information and Digital Technologies (IDT), 2016, pp. 261-265. IEEE

[23] B. Zhang, C. Yin, B. David, Z. Xiong and W. Niu, "Facilitating professionals' work-based learning with context-aware mobile system." Science of Computer Programming, 2016, vol. 129, pp.3-19

[24] A. Zarrad, and A. Zaguia, A.,"Building a dynamic context aware approach for smart e-learning system." In 2015 Second International Conference on Computing Technology and Information Management (ICCTIM), 2015, pp. 144-149. IEEE

[25] J. Brank,M. Grobelnik and D. Mladenic, D., "A survey of ontology evaluation techniques." In Proceedings of the conference on data mining and data warehouses (SiKDD 2005), 2005, pp. 166-170). Citeseer Ljubljana, Slovenia.

The Impact Of Quantum Genetic Algorithms In Minimizing Task Migration' Overheads

***Abstract*—Genetic algorithms are widely used to solve problems in complex system for its easiest implementation and achievability of global optimum with a proper approximate solution, but quantum-inspired evolutionary algorithms have been surpassing the classical algorithms for their abilities to solve problems of polynomial-time that is considered it as impossible to be solved but with million years while their advantage is to balance between exploration and exploitation of the solution space and also obtain better solutions, even with a small population. In this paper, we propose a novel task migration algorithm that minimize the overall overheads caused during the migration process in network on chip (NoC) while searching the appropriate checkpoints during run-time using quantum genetic algorithm.**

I. INTRODUCTION

NoC technology is often called "a front-end solution to a back-end problem." [1]. The more critical the NoC becomes, the more reliable it needs to be [2][3]. In modern application embedded systems, the performance constraints need to be satisfied according to the number of computational elements supported by the system. NoC is the high performance and scalable alternative to the system on chip architecture [3]. However, such systems are usually operated by self-source power like batteries, so to increase the operating time, a minimization of the energy consumption during their design is required. Due to the heterogeneity of the architecture, assigning the same task to different processing elements (PEs) leads to very different energy consumption values associated to computation. The application mapping that determines the assigning of multimedia application tasks to NoC tiles in such a way that the energy consumption and the latencies are optimized is considered as non-deterministic polynomial-time hard (NP-hard) problem where the search space of the problem increases factorially with the system size and due to that an effective optimization algorithm is required. The choice of the most appropriate algorithm is a critical issue because an optimal mapping may enhance NoC performance up to 60%. For that many heuristic algorithms are designed to find a near optimal mapping such as genetic algorithm. Also, the classical chromosomes are weak in representing the population diversity whereas the use of quantum bit representation leads to better population diversity and also the use of quantum gate drive the population towards the best solution. In this paper, we propose a novel task migration algorithm that minimize the overall overheads caused during the migration process while searching the appropriate checkpoints during run-time using quantum genetic algorithm.

II. RELATED WORK

The scheduling problem is a traditional research topic, almost all previous work focuses on maximizing the performance through the scheduling process [7]. The algorithms developed this way are thus not suitable for real-time embedded applications where the objective is to minimize the energy consumption under tight performance constraints [8]. In [9], two mapping algorithms, one based on simulated annealing and one based on genetic algorithm for energy- and communication-aware mapping problems of mesh-based NoC architectures are proposed. In [10], a novel reliability-aware hard real-time task scheduling method for multicore systems along with a quantitative reliability model was performed.

In [11], the author presented a low time-complexity heuristic mapping algorithm of weighted application graph under permissible bandwidth constraint to minimize communication energy of 2-D mesh-based NoC architecture. In [12], the author proposed a self-optimizing and self-programming computing system design framework that achieves both programmability and flexibility and exploits computing heterogeneity. In [4],[5], a comparison is made between a classical genetic algorithm and a quantum inspired method for the travelling salesperson. An improved QGA based on multi-qubit encoding and dynamically adjusting the rotation angle mechanism was presented to separate the blind sources [6].

Many task migration mechanisms were developed such as: checkpoints, debug registers, incoming migrations, accepting task migrations by polling, and accepting task migrations by interrupts. The choice of migration mechanisms is based on the NoC platform used or type of applications designed on that NoC architecture. A migration checkpoint is a physical point in the program where a task migration is possible after the monitor decision. The point is often provided manually by the programmer and should be put in an adequate location. the responsibility of choosing the checkpoints is hereby put on the programmer. In this paper, we will use the migration checkpoints due to the random property of quantum genetic algorithm, that is organized as follows: in Section 3, task migration mechanisms are discussed, followed by Section 4 where the used preliminaries are presented. Section 5 is dedicated to describe the overheads minimization using quantum genetic algorithm and Section 6 concludes this paper.

III. TASK MIGRATION MECHANISMS

A task migration model is a way of showing what actually happens during a migration on a higher level of abstraction. The model shows which actions need to be taken, and what platforms support such actions. Assuming a migration is taking place between two CPU cores. The target core must have all the necessary information needed for resuming the migrated task. In this paper, we classified that information based on the migration difficulty (information difficult to migrate and information easy to migrate) because the heavy overhead of the communication delay or moving the tasks physically is one of the main concerns during the migration. The first class contains the information that have a little impact on the total overhead such as: *CPU state, task*

stack, and cache memory. Whereas the second class contains most of information that have a huge impact on augmenting the energy or the delay overheads such as: *Heap data, Global variables, program code, and task associations*.

A. CPU state

A task needs to know its CPU state. The CPU state is easily read from the CPU status register and is stored in a temporary variable.

B. Task stack

The task stack holds the variables used in the task and the values of these. The task stack is fairly easy to migrate because the system knows where the stack begins and its size.

C. Heap data

Data allocated in the heap is difficult to migrate since the whole core uses the same heap. Variables associated with the task must be somehow marked to inform the migration mechanism which data to migrate. Problems could eventually occur when determining the size of dynamically allocated arrays in the heap. Assuming there is a mechanism for linking heap data to tasks, the data could be migrated and re-allocated on the other core.

D. Program code

Another difficult part to migrate is the executing code in the program memory. The program code is statically stored, and must also be marked with associations to the relevant task. The program counter in the system must continue exactly from the migration point relative to the program. The migration of the program code is also substantive when considering the run-time update.

E. Task associations

Various associations exist between tasks, such as information exchange, calls, child tasks etc. A task is able to communicate with other tasks using message queues [13], which can only be used if the tasks are executing on the same core. Otherwise the interface of communication must change, so that the tasks communicate over some kind of inter-core communication.

F. Cache memory

To speed up cache searches and to eliminate the cold start of cache memory, possible to migrate the whole cache memory or cache areas is needed.

G. Global variables

Several tasks could use the same global variables, which leads to a synchronization problem. For security reasons tasks that are selected for migration should preferably not share global variables with other tasks, since the communication using global variables could be broken or slowed down if one of the tasks is moved to another core.

IV. PRELIMINARIES

A. Platform Characterization

The NoC main components are depicted in Fig.1. A typical service is a bidirectional communication channel that transfer packets which are split into several flits. The node of the network is composed of PE and a router. The router is connected to the four neighboring tiles and its local processing core via channels. ***Formally***: $G(P, L)$a directed topology graph where P_i represents a node of the network, a directed arc $L_{i,j} = (P_i, P_j)$ represents a physical unidirectional channel connecting two nodes P_i and P_j.

B. Application Model

A multimedia application consists of tasks and channels. ***Formally:*** the application is given by; $G(\mathrm{C}, \mathrm{E})$: is a synchronous dataflow graph, with each vertex $c_i \in \mathrm{C}$ represent a task defined by its $d(i)$ deadline, $pr(i)$ its priority level and $w(i)$ the task workload. The directed ed $e_{i,j} \in \mathrm{E}$ represent the communication between the tasks c_i to c_j.The weight of edge $v(c_{i,j})$ represents the bandwidth required of the communication from c_i to c_j.

C. Power Model:

The power model used to estimate the energy consumption is given by:

$$E_{Noc} = N_{task} \times E_{c_i}^{P_j} + N_{channel} \times E_{e_{i,j}} \quad (1)$$

where: N_{task} is the number of tasks, $N_{channel}$ is the number of channels, $E_{c_i}^{P_j}$ defined as the number of cycles it takes to execute c_i on processor p_j:

$$E_{c_i}^{P_j} = r_j^i \times f_p \times E_{c_{ip}}^{p_j} \quad (2)$$

r_j^i is the execution time of c_i on p_j, f_p is the operating frequency. $E_{c_{ip}}^{p_j}$ is energy of one computational interval. The energy consumed due to sending the message $V(\mathrm{c}_{i,j})$:

$$E_{e_{i,j}} = NB_F \times \frac{v(c_{i,j})}{f\_z} \times E_{flit} \quad (3)$$

E_{flit} is the energy consumed during sending one flit, f_z is the flit size, NB_F is number of flits.

D. Delay model:

The time it takes to execute and transmit all flits is given by:

$$L_{Noc} = \sum_{i=1}^{T} Ex_{c_i}^{p_j} + \sum_{j=1}^{C} (n-1) \times Com_j \quad (4)$$

The execution time $Ex_{c_i}^{p_j}$ defined as the number of cycles it takes to execute c_i on p_j and it is computed as follows:

$$Ex_{c_i}^{p_j} = \frac{w(i)}{f} \quad (5)$$

The communication time is the time it takes to send the flits through the link and it is computed as follows:

$$Com_j = \frac{N_{flit}}{bw} \quad (6)$$

E. *Problem formulation*

Mapping and schedule strategy of the application graph $G(C,E)$ and the NoC graph $G(P,L)$ defined by: $Map: C \to P, s.t. map(c_i) = p_i, \forall c_i \in C, \exists p_i \in P$

$schedule: T \to T, s.t. schedule(\forall c_i \in C\ mapped_{to}(p_i))$

$minimise\ (E_{Noc}, L_{Noc})$

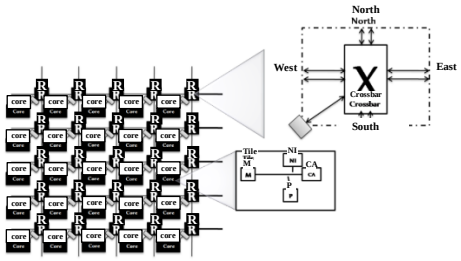


Fig. 1. NoC- based system.

V. Overheads minimization using quantum genetic algorithm

The solving of the aforementioned problems is performed with the following five steps during the scheduling:

A. ***Step 1***. *Real-Time Scheduling Analysis*

We assume that each core has its own priority-ordered queue to schedule tasks, for that five parameters are considered for each task, C_{i_W} is the task workload which is amount of job of each task and C_{i_D} is the task deadline that is the moment of time where the task to should finish its execution(for soft real-time), and C_{i_F} is the first one reached its deadline that is the nearest workload to the deadline, and C_{i_S} is the minimum slack that is task deadline minus the task workload with

considering the actual time and finally , $C_{i_{Avr}}$ is the average response time that defined in Eq.7, where $H_{priority}(c_i)$: denotes the set of higher priority tasks running on the same core as task c_i.

$$r_i^{n+1} = C_{i_w} + \sum_{\forall c_j \in H_{priority}(c_i)} \left(\frac{r_i^n}{period}\right) \times C_{j_w} \qquad (7)$$

B. ***Step 2.*** *Elimination of buffer wait time and the task wait time (The mapping strategy)*

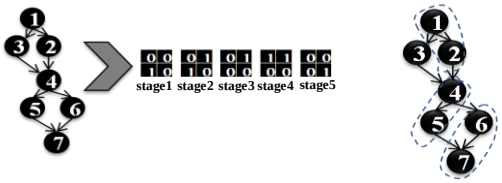
The main idea of this technique is to map the tasks with higher priority onto different processors in which mean the tasks that are executed in a parallel way are allocated each one to a different processor. Seeing Fig.2 (a), the application has 7 tasks implemented onto 2x2 NoC, the number of stages is 5, the first contains task number 1, stage2 contains task 2 and 3 till stage 5 that contains task 7. In this technique we have two cases; the first is when the number of tasks per stage is less than the number of processors, so we map the tasks randomly where each one occupies a different processor according to Eq.8. The second is the opposite of the first; where the number of tasks per stage is bigger than the number of processors, so we sort the tasks according to the scheduling strategy used in *step1*, and we map them on the processor that contains the task with higher priority as it is described with Eq.9.

$$\forall_{stage_i} \in STAGE / c_j \in C; \exists p_k \in P / p_k \subset stage_i$$

$$if\ P_k = free, c_j \to p_k \qquad (8)$$

$$if\ P_k \neq free, min = sort\left(w_{(j \in stage_i)}\right), c_i \to p_{min} \qquad (9)$$

After applying *Step2*, we the elimination of the buffer wait time and the task wait time in case of two tasks are allocated to the same processor is accomplished in despite of the scheduling strategy used in *Step1*.



(a) An example of the Step2. (b) An example of Step3.

Fig. 2. The mapping strategy.

C. ***Step 3***. *Association mapping*

Another metric is used between each two stages, we call it the "Association mapping" that maps two linked tasks to the same processor and it will look like the entire arc with its sink and source tasks is allocated to that processor. A random mapping is used to map the other cores that have not assigned to any processor (such as core3 in Fig. 2 (b)) using (10) or (11).

D. ***Step 4***. *Task migration model*

The migration process is done in two parts, the first one is before the migration where the migration decision will be allowed or not. The second part is on the task run-time while assuming a migration is taking place between two CPU cores.

***The first part*:** is done by the load balancing monitor or as we call it, the migration decider. This decision is taken based on the CPU occupation, an energy policy, and all the performances that are need to be satisfied, while taking into consideration all the trade-offs and overheads provided by the nodes. The decision is allowed if the following condition is satisfied:

$$\sum_{k=0}^{M} P_j(C_{k_{ext}}) > \sum_{k=0}^{M} P_j(C_{k_{deadline}}) \qquad (10)$$

Eq.10 indicates that if the sum of execution time of tasks mapped to the same processor leads to missing their deadlines, then the load balancing monitor enable the task migration.

The second part: is usually started after the checkpoints selection in the program, the operating system stops the currently running task by suspending it. The data associated with the task is stored and moved onto the target core by the migration mechanism. The operating system can, if the migration was successful, manually switch in a new task for processing.

The contribution in this paper is to minimize the overall overheard during the task migration while considering only the second class that contains most of information that have a huge impact on augmenting the overheads of the energy consumption and the latencies of the systems, while the other information (first class) are negligent.

Formally, given task C_i that started the execution on processor P_{j1},and migrated on processor P_{j2}, $HD_{ij_1j_2}$, $GV_{ij_1j_2}$, $PC_{ij_1j_2}$, $and\ TA_{ij_1j_2}$, are the relative information heap data, global variables, program code, and task associations respectively of task C_i to and the overall overheads are the sum of energy consumption (E_{OH}) and latency (T_{OH}) of migrating those information, and it is calculated as follows :

$$mig_{OH} = \sum_0^D T_{OH}(HD_{ij_1j_2} + GV_{ij_1j_2} + PC_{ij_1j_2} + TA_{ij_1j_2}) + E_{OH}(HD_{ij_1j_2} + GV_{ij_1j_2} + PC_{ij_1j_2} + TA_{ij_1j_2}) \quad (11)$$

- ***Remark**: The overheads are calculated with considering the energy and latencies during transferring the associate information to the target processor P_{j2} and also the de-allocation (deleting) of those information from the old processor P_{j1}, without considering the difference between the simulated values, in which mean, we will not remove from the overheads the energy consumption and latencies values of the current task in the case of that task did not migrated, because the decision of migration will be taken during run-time, and it will be impossible for us to predict if two tasks will wait for each other or will be migrated.*

1) ***Overhead delay:*** represents the time it takes for the task to be migrated from P_{j1} to P_{j2}, and it is given by:

$$T_{OH} = \sum_{j=1}^{c}\left[(n_r - 1) \times [wl_{ci} + transfer_i] + n_r \times wb_j\right] \quad (12)$$

- wl_{ci} is the link wait time of task C_i is the time of waiting into the buffer in case of non-availability of channels (or virtual channels).
- n_r is the number of hops traversed.
- wb_j is the time it takes for the deleting the information for task C_i from the buffer.
- The transfer time is the time it takes to send the task C_i through the link and it is computed as follows:

$$transfer_i = \frac{size\_c_i}{bw} \quad (13)$$

2) ***Overhead energy consumption:*** the energy consumed due to migration the task C_i is calculated as follows:

$$E_{OH} = n_r \times E_{router} + (n_r - 1) \times E_{link} \quad (14)$$

E_{router}, E_{link} are the energy consumed during the migration from the router and the link respectively. The parameters values are obtained from the Predictive Technology Model (PTM [14]), and [2][15][16].

In order to calculate the energy consumed in an interconnect wire, a study of its physical characteristic, and its electric behavior has been done [15] [16] [17]. As the wire got longer, repeaters were used to minimize the energy consumption. The power equation for one gate driven wire is given by:

$$P_{link} = \frac{1}{2}CV^2\alpha_l fV + \tau\alpha VI_{short}f + VI_{bias,wire} + VI_{leak,gate} \quad (15)$$

C is the load capacitance. V is the supply voltage, α_l is the switching activity of the gate and f is the operating frequency of the system.

$\tau\alpha VI_{short}f$ is the short time period during I_{short} flow, and $I_{bias,wire}$ and $I_{leak,gate}$ are the drain and source to body junction leakage currents in the NMOS device. The total energy consumption in a router during the migration is given as follows:

$$E_{router} = E_{FIFO} + E_{Arbiter} + E_{Crossbar} \quad (16)$$

$E_{Arbiter}$ is the energy consumed due to setting up of a path for c_i to traverse from the input port to the output, and it is calculated as follows: ($E_{int_a} = 0.014\alpha + 0.0084, E_{swit_a} = 0.8415\alpha + 0.45, E_{leak_a} = 0.0047$).

$$E_{Arbiter} = E_{int_a} + E_{swit_a} + E_{leak_a} \quad (17)$$

$E_{Crossbar}$ Energy consumed due to routing the packets from input ports to the output, and it is calculated as follows:($E_{int_c} = 1.002\alpha + 0.515, E_{swit_c} = 1.02348\alpha + 0.1465,\ E_{leak_c} = 0.005$)

$$E_{Crossbar} = E_{int_c} + E_{swit_c} + E_{leak_c} \quad (18)$$

E_{FIFO}: Energy due to read only (the deleting of task c_i):

$$E_{FIFO} = P_{read} + P_{clk} + P_{int} \quad (19)$$

$$P_{read} = \mathrm{r_r} \times \mathrm{P_{control}} + \alpha \times \mathrm{P_{retrieve}} \quad (20)$$

Here, r_r is the rate at which read actions occur, $P_{control}$, $P_{retrieve}$ is the power consumed at the control unit and the power consumed to retrieve the data from the FIFO respectively.

P_{clk} Clock activity causes permanent constant power consumption due to switching activity.

P_{int} Which is the result of the internal short circuits occurring during switching of bits in the incoming data, the switching of read actions and clock activity. Therefore P_{int} can be calculated as follows:

$$P_{int} = rk_1 + k_3 \quad (21)$$

Here, r is the rate of read actions, the constants k_1 is the average internal power consumed for the control of read actions and due to bit changes in data and k_3 is due to the clock activity.

E. ***Step 5.*** *Minimizing the overheads searching the appropriate checkpoints using QGA:* This step plays an important part in our algorithm, in following, a detailed mechanism of how using QGA to minimize the overall overhead, with the system performances.

1) Problem Representation

In QGA, we used the same chromosome representation as CGA but in place of bit, a Q-bit is used. Q-bit is defined as the smallest unit of information. A Q-bit can be represented as:

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \quad (22)$$

Where α and β are real numbers whose represent the probabilities that the Q-bit found in the "0" and "1" states, respectively with $|\alpha|^2 + |\beta|^2 = 1$.

The state of a Q-bit may be "0", "1", or a linear superposition of the two:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (23)$$

Where $|0\rangle$ and $|1\rangle$ mean the states "0" and "1", respectively. However, a linear superposition of states with m Q-bits can be represented by Q-bit individual as it given in Eq.7. For that, a system with m Q-bits can represent 2^m states at the same time (for $i = 1,2,\dots m$: $|\alpha_i|^2 + |\beta_i|^2 = 1$).

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_m \\ \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_m \end{bmatrix} \quad (24)$$

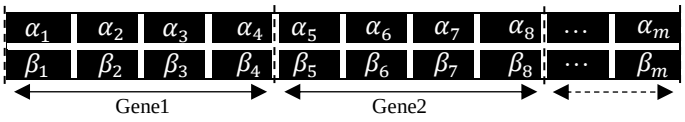


Fig. 3. Quantum chromosome structure

2) Generation of the initial population: The population is the main element of a quantum genetic algorithm, where its size influences the coverage of mapping space. A population that is too large takes time to evolve whereas a population that is too small lead to a local minimum. We generated an initial population, $P = 100$ individuals are chosen randomly and well distributed in the research space.

First, the generation counter is set to $t = 0$, then an initialization of the group of Q-bit individuals $Q(t)$ where $Q(t) = [q_1^t, q_2^t, \dots, q_n^t]$, n is the total number of Q-bit individuals and q_j^t is the j_{th} Q-bit individual at generation t which is defined as: (where $j = 1,2,\dots n$, m is the string length)

$$q_j^t = \begin{bmatrix} \alpha_{j1}^t & \alpha_{j2}^t & \alpha_{j3}^t & \cdots & \alpha_{jm}^t \\ \beta_{j1}^t & \beta_{j2}^t & \beta_{j3}^t & \cdots & \beta_{jm}^t \end{bmatrix} \quad (25)$$

*3) Measuring Chromosomes (Generate X(t) by measuring of Q(t)):*In this step, each Q-bit is observed and measured from $Q(t)$ in order to extract a classic chromosome $X(t)$ that the evaluation of each quantum chromosome. For that, $X(t)$ is a group of binary solutions where $X(t) = [X_1^t, X_2^t, \dots, X_n^t]$,and $x_{j1}^t = [x_{j1}^t, x_{j2}^t, \dots, x_{jn}^t]$,where x_{j1}^t is binary value that is determined as:

$x_{ji}^t \coloneqq get\ x_{ji}^t\ in\ [0,1]$;

$if(x_{ji}^t < |\beta_{ji}^t|^2) x_{ji}^t = 1; else\ x_{ji}^t = 0$;

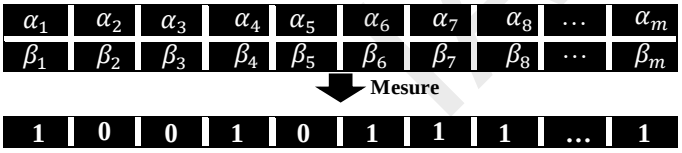
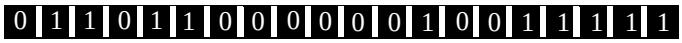


Fig. 4. Measured chromosome.

A migration checkpoint is a physical point in the program where the migration is allowed with the monitor decision. The point is often provided manually by the programmer and should be put in an adequate location. the responsibility of choosing the checkpoints is hereby put on the programmer, in this paper we defined four (04) checkpoints for each task. For that, we chromosome encoding will include and charged of allowing or denying the migration on predefined checkpoints. An example is shown in Fig.5, given 4 checkpoints, and 4 tasks that are mapped and scheduled based on the aforementioned steps. After generating the chromosome (see Fig.5 (a)), it will be divided on the number of tasks, and each subdivision contains 5 genes, the size of the chromosome is 5×Nb_task. An example of the first subdivision is described as follows: the first 4 genes represent the 4 checkpoints corresponding to the 1st task, the 5th gene represents the monitor decision. The 1 and 0 in the monitor decision mean that the migration is allowed or denied respectively. In case of the migration is not allowed, the program won't check if there is 1 on the other column such in Ex2, the monitor decided to deny the migration, for that the program wont check if there is 1 on others such in Ex1, and it will be adjusted and modified to be 0 (see Fig.5 (c)). Generally, for a set of task C, the chromosome encoding format after the mapping and the scheduling on the appropriate NoC architecture is illustrated in Fig.6.



(a) Measured chromosome.

	1st checkpoint	2nd checkpoint	3rd checkpoint	4th checkpoint	Monitor decision
Task 1	0	1	1	0	1
Task 2	1	0	0	0	0
Task 3	0	0	1 (ex1)	0	0 (ex2)
Task 4	1	1	1	1	1

(b) Chromosome subdivision.

	1st checkpoint	2nd checkpoint	3rd checkpoint	4th checkpoint	Monitor decision
Task 1	0	1	1	0	1
Task 2	1	0	0	0	0
Task 3	0	0	0	0	0
Task 4	1	1	1	1	1

(c) Structure of the chromosome after adjustment phase.

Fig. 5. An example of the problem formultion (structure of the chromosome).

	1st checkpoint	2nd checkpoint	3rd checkpoint	4th checkpoint	Monitor decision
Task 1	0	1	1	0	1
Task 2	1	0	0	0	0
….	….	….	….	….	….
Task C	1	1	1	1	1

Fig. 6. The general structure of chromosomes.

4) Fitness function: represents the desired optimization goal. For that, it is multi-objective optimization and a combination of energy consumption, latencies, and overheads. The evaluation of each classical chromosome $X(t)$ is through its evolution throughout K iterations. A note between [0,1] is assigned to individuals. The fitness function is given by:

$$Fitness_{Noc} = \frac{\lambda_1 \times E_{Noc} + (1-\lambda_1) \times E_{OH} + \lambda_2 \times L_{Noc} + (1-\lambda_2) T_{OH}}{Fitness_{min}} \quad (26)$$

$Fitness_{min}$ is the minimum energy, λ_1, λ_2 are proportionalities coefficients which are used to adjust the

proportion of energy consumption and latencies in cost function, and the value range is $0 < \lambda_1 + \lambda_2 < 1$, with $\lambda_1 = \lambda_2$, we choose to equal the energies ($\lambda 1$) and the latencies ($\lambda 2$).

5) Store the best solution among $X(t)$ into $B(t)$: $B(t)$ is a matrix that stores the best solution in the whole population.

6) Update $Q(t)$ using Q-gates: Quantum individuals are updated by using Q-gates. The population $Q(t)$ is updated with a quantum gates rotation of qubits constituting individuals. The rotation gate $U(\Delta\theta_i)$ and the update operation are expressed as:

$$U(\Delta\theta_i) = \begin{bmatrix} \cos(\Delta\theta_i) & -\sin(\Delta\theta_i) \\ sin\ (\Delta\theta_i) & \cos(\Delta\theta_i) \end{bmatrix} \quad (27)$$

$$\begin{bmatrix} \alpha_{ji}^t \\ \beta_{ji}^t \end{bmatrix} = U(\Delta\theta_i) \begin{bmatrix} \alpha_{ji}^{t-1} \\ \beta_{ji}^{t-1} \end{bmatrix} \quad (28)$$

$\Delta\theta_i$ is a rotation angle which determines the magnitude and direction of rotation.

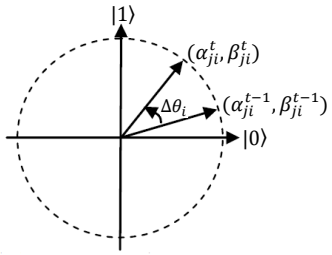


Fig. 7. Polar plot of the rotation gate for Q-bit individuals.

TABLE I. LOOK-UP TABLE FOR QUANTUM GATES ROTATION.

x_{ji}^t	b_i^t	$f(X_j^t) > f(B^t)$	$\Delta\theta_i$	$S(\alpha_{ji}^t \times \beta_{ji}^t)$		
				>0	<0	$=0$
0	0	0	$\Delta\theta_2$	-	+	±
0	0	1	$\Delta\theta_2$	-	+	±
0	1	0	$\Delta\theta_1$	-	+	±
0	1	1	$\Delta\theta_2$	-	+	±
1	0	0	$\Delta\theta_1$	+	-	±
1	0	1	$\Delta\theta_2$	+	-	±
1	1	0	$\Delta\theta_2$	+	-	±
1	1	1	$\Delta\theta_2$	+	-	±

Fig.7 illustrates the polar plot of the rotation gate for Q-bit individuals. At generation t, the rotation angle $\Delta\theta_i$ is updated according to the criteria summarized in Table.1, where x_{ji}^t and b_i^t are the binary control variables in solution X_j^t and the best solution B^t of $B(t)$, respectively. $f(X_j^t)$ and $f(B^t)$ represent the objective function values of X_j^t and B^t. For example, when x_{ji}^t and b_i^t are 0 and 1, and $f(X_j^t)$ is larger than $f(B^t)$, the rotation angle $\Delta\theta_i$ is updated according to $S(\alpha_{ji}^t \times \beta_{ji}^t)$ in Table.1 where $S(\alpha_{ji}^t \times \beta_{ji}^t)$ is the sign of $\alpha_{ji}^t \times \beta_{ji}^t$.

In the last step, the best solution among $X(t)$and $B(t-1)$is stored to $B(t)$, and terminated if the stopping conditions are met; else generate a new population.

VI. CONCLUSION

In this paper, we benefit with the use of quantum genetic algorithms to optimize NoC performances, of better population diversity with the use of quantum bit, and a strong orientation to best solutions with the use of quantum gate. we proposed a novel task migration algorithm that minimize the overall overheads on both energy consumption and latencies caused during the migration process in NoC through five steps while searching the appropriate checkpoints during run-time using.

REFERENCES

[1] Wooseok Lee, "ArterisFlexNoCResilience Packagel", Design Automation for Embedded Systems Journal (2014).

[2] S. Kumar, A. Jantsch, J.-P.Soininen, M. Forsell, M. Millberg, J. Oberg, K. Tiensyrja, and A. Hermani, ―A New on Chip Architecture and Design Methodology,‖ in Proc. IEEE Computer Society Annual Symposium on VLSI, Apr. 25-26, pp. 105-112 (2002).

[3] L. Benini, G. De Micheli, Networks on chips: a new soc paradigm. Computer vol. 35, no.1, pp. 70–78 (2002).

[4] Ajit Narayanan and Mark Moore. Quantum-inspired genetic algorithms. In Evolutionary Computation, 1996., proceedings of IEEE International Conference on, pages 61–66. IEEE, 1996.

[5] Z. Laboudi and S. Chikhi, "Comparison of genetic algorithm and quantum genetic algorithm. The international Arab Journal of Information Technology", vol. 9, no.243, (2012).

[6] Zhenquan Zhuang Junan Yang, Bin Li. Research of quantum genetic algorithm and its application in blind source separation. Journal of Electronics, vol. 20, no.1, pp. 62–68 (2003).

[7] S. Kumar, "A two-step genetic algorithm for mapping task graphs to a network on chip architecture", Proc. Euromicro Symp. Digit. Syst. Des., pp. 180-187 (2003).

[8] P. K. Sahu, S. Chattopadhyay, "A survey on application mapping strategies for Network-on-Chip design", J. Syst. Architecture, vol. 59, no. 1, pp. 60-76, (2013).

[9] S. Tosun, O. Ozturk, E. Ozkan, M. Ozen Application mapping algorithms for mesh-based network-on-chip architectures J. Supercomput, vol. 71, no. 3, pp. 995-1017 (2015).

[10] N. Alireza, S. Saeed, M. Siamak, "CMV: Clustered Majority Voting Reliability-Aware Task Scheduling for Multicore Real-Time Systems" IEEE Transactions on Reliability, Vol. 68, no.1, pp. 187 – 200 (2019).

[11] Pradeep S. Kumar, B. Santosh, M. Pinaki, "Energy efficient heuristic application mapping for 2-D mesh-based network-on-chip", Microprocessors and Microsystems, Vol.64, pp. 88-100 (2019).

[12] X. Yao, N. Shahin, B. Paul, Self-Optimizing and Self-Programming Computing Systems: A Combined Compiler, Complex Networks, and Machine Learning Approach, IEEE Transactions on Very Large Scale Integration (VLSI) Systems, pp.1-12 (2019).

[13] Seo Y, Kim J, Seo E, ―Effective analysis of DVFS and DPM in mobile devices,‖ Journal of computer Science and Technology ,Vol. 27, no 4, pp. 781–90 (2012).

[14] E. Rijpkema, K.G.W. Goosens, A.Radulescu, J.Dielissen, J.van Meerbergen,P.Wielage, and E.Waterland,―Trade Offs in the design of a Router with Both Guaranteed and Best- Effort Services for Network on Chip,‖ in Proc. Design, Automataion and Test in Europe Conference and Exhibition, pp. 350-355 (2003).

[15] R. Mulllins, A. West, and S. Moore, ―Low-Latency Virtual-Channel Router for On-Chip Networks,‖ in Proc. 31st Annual International Symposium on Computer Architecture, vol.19, no.23,pp.188-197 (2004).

[16] A.Mello, L. Tedesco, N. Calazans, and F. Moraes, ―Virtual Channels in Network on Chip: Implementation and Evaluation on Hermes NoC,‖ in Proc.th Symposium on Integrated Circuits and Systems Design, pp. 178-183 (2005).

[17] J. William, Dally and brian towles. route packets, not wires: on-chip interconnection networks, in Proceedings of the Design Automation Conference, pp. 684–689 (2001).

Facial Expression Recognition System

Chebah Wafa
Department of Infomatics
Badji Mokhtar University
Annaba-B.P.12, Annaba, 23000 Algeria
ouafa.chebah@univ-annaba.org

Laskri Mohamed Tayeb
Department of Infomatics
Badji Mokhtar University
Annaba-B.P.12, Annaba, 23000 Algeria
laskri@univ-annaba.org

***Abstract*— One of the non-verbal communication method by which one understands the mood state of a person is the expression of face. Automatic facial expression recognition (FER) has become an interesting and challenging area for the computer vision field and its application areas are not limited to mental state identification, security, and automatic counseling systems, face expression synthesis, lay detection, music for mood, automated tutoring systems, operator fatigue detection. In this paper, we propose a simple approach to recognize facial expressions from static images. First, the noise-removal/enhancement is done in the preprocessing step by taking image as an input and gives the face for further processing, then the facial component detection detects the ROI for eyes, nose, mouth, eyebrow. The feature extraction step deals with the extraction of features from the ROIs. These features are used to obtain the corner point features, which will be used to feed as feature input vectors to the multilayer perceptrons (MLPs). The MLPs will hence learn from these inputs to make classification decisions of the final output emotion. Experimental results demonstrate an average recognition accuracy of 74 % in the FABO database [1].**

***Keywords*— facial expressions, human face, emotional recognition, neural network, multilayer perceptrons .**

I. INTRODUCTION

Facial Expression is one of the most powerful, nature, and immediate means for human beings to communicate their emotions and intentions. Mehrabian [2] shows the fact that 55% of the emotional message is communicated by the facial expression whereas only 7% by the language and 38% by the paralanguage. Thus, facial expressions play a predominant role in terms of coordination and have a greater impact on the listener than the textual content of the expressed message.

An automatic facial expression recognition system involves two crucial parts: facial feature representation and classifier design. Facial feature representation is to extract a set of appropriate features from original face images for describing faces. Mainly two types of approaches to extract facial features are found: geometry-based methods and appearance-based methods. In the geometric feature extraction system, the shape and location of various face components are considered. The geometry-based methods require accurate and reliable facial feature detection, which is difficult to achieve in real time applications. In contrast, the appearance-based methods, image filters are applied to either the whole face image known as holistic feature or some specific region of the face image known as local feature to extract the appearance change in the face image. Classifier design classifies the facial expressions based on extracted relevant features. Major classifier include: template matching, support vector machine (SVM), KNN, neural networks, HMMs, etc.

II. RELATED WORKS

Automatic facial expression recognition involves three vital aspects: face detection, features extraction, classification and recognition of facial expression

There are a number of methods available for face detection/extraction but Viola-Jones [4] is the most prominent. It is for the object detection and it can detect different body parts like face, upper body part, head etc. as well as facial components like nose, mouth, eyes etc

Many studies have been done on facial features extraction, in [5] Black and al. used local parametric models to represent the movement of faces. They estimate the relative movement of the characteristic features in the reference of the face; in [6] Cohn et al. proposed a hierarchical algorithm to track characteristic features by estimating optical flow. Displacement vectors represent information about changes in facial expression. Similarly, In [7] Padgett and al. realized eye and mouth templates, calculated by Principal Component Analysis of a learning set, in conjunction with neural networks. On the other hand, in [8] Hong et al. proposed a global model based on labeled graphs constructed from landmarks distributed on the face.. In [9] Cootes et al. used an active appearance model representation (AAM) to automatically extract parameters characterizing the face.

The literature presents a significant number of techniques that deal with the recognition of facial expressions, in [10] Huang and al. have seen that the study of 44 units of action (AUs) is not easy and they proposed to replace them with 10 parameters of actions (APs). They removed the 10 APs based on the difference between a neutral face and another expressive and applied a PCA on them to reduce the dimension to two and simplify the recognition process. The elaborate system succeeded in recognizing the expressions studied and had a recognition rate equal to 84.5% [11].The works [12] and [13] of Lyons et al. presented a system for classifying face images into categories according to three criteria: sex, race and facial expression. The recognition rate of the system reaches 92% when tested with 193 images of nine Japanese women and decreases to 75% when tested with unfamiliar subjects [11] [12]. In [14], Valstar et al. proposed a fusion approach between different levels of abstraction of types of facial features, Zhang

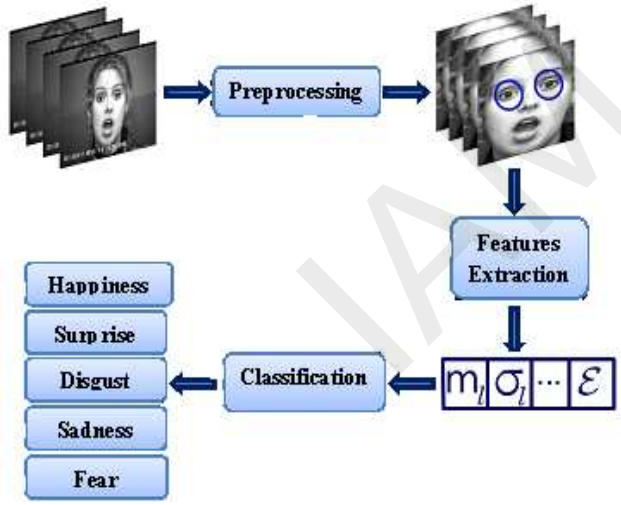
et al. [15] used neural networks for the facial expression recognition problem. They marked on the face 34 points called points of interest and calculated 18 Gabor wavelet coefficients at each point to make a set of characteristics geometric position type and Gabor type filter. The recognition performance of the seven expressions by the system breaches 90.1%. In [16], Cohen et al. presented a multi-level HMMs method for segmentation and facial expression recognition from a video clip.The first level of the proposed architecture is composed of six independent HMMs, one HMM for each of the universal expressions. Franck Davoine et al. [17] proposed a solution that recognizes the expression of a face by linear discrimination. For classification, the authors used the Mahalanobis distance as a criterion of similarity between the MAA vector of the test image and the mean vector in Fisher's space. In [18], Hammal and al. use the theory of evidence based on either permanent traits or transient features of the face and average 64.5% for recognition of (joy. or disgust) and (surprised or afraid).

III. Facial Expression Recognition System

The objective of the Facial Expression Recognition System (FERS) is the extraction of facial features from faces and classifies emotions in one of the six universal emotions: happiness, sadness, anger, disgust, surprise and fear.

The Facial Expression Recognition System (FERS) can be divided into three blocks performing: face detection, features extraction, and sample classification, the following Fig.1 shows the FERS architecture and its various stages.

Fig. 1. General architecture of FERS



First, the captured images are preprocessed and faces are detected. Next, we locate the Region Of Interest (ROI) for eyes, eyebrow, mouth and nose. From the detection windows obtained for the different facial component, we use these RoI windows to obtain the corner point features. The feature points extracted represent inputs for the neural network. On the basis of these input feature vectors, the neural network will hence learn to make classification decisions of the final output emotion.

A. Face detection

There are various methods for the face detection described in the literature. These methods present a difference in their complexity, performance, type, and nature of the images. In this work we chose a face detection algorithm based on the skin color by using RGB space [19] because it offers good performance (Fig.2.b).

A pixel is said a skin pixel, if the components R (Red), G (Green) and B (Blue) of the color satisfy the following conditions (segmentation and binarization):

If (R > 95) and (G > 40) and (B > 20) and ((Max [R, G, B] – Min [R, G, B]) > 15) and Abs (R − G > and R > G and R > B) then skin pixel 1 else skin pixel 0

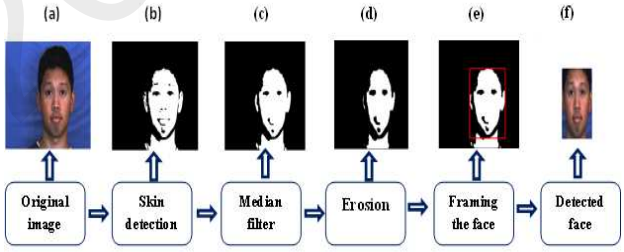
After selecting the skin pixels, we obtain a binary image whose various pixel values are equal to 0 or 1. Then we improve the image obtained by median filter (Fig.2.d) and mathematical morphology (erosion).

The principle of determining the coordinates of the face is to achieve a horizontal projection (HPI) and a vertical one (VPI) of the binary image.

$$HPI = \Sigma^{m}_{y=1} I(x,y) \qquad (1)$$

$$VPI = \Sigma^{n}_{x=1} I(x,y) \qquad (2)$$

Fig. 2. Detection and extraction of the face



B. Facial features extraction

Several extraction methods use the color information [20]. These methods have two limitations. They work only with color images, so illumination changes can affect the results. There are some methods based on the gradient information that are robust to illumination changes. We chose a method developed in [21]. The method is based on the fact that human faces are constructed in the same geometrical configuration. To model the face with a geometric model, the method uses a gradient tracking to locate each component of a frontal face. Thus, it presents robustness against small rotation angles, lighting conditions, skin color or accessories (mustache, glasses, beards) [21].

B.1) Facial features localization

For facial features localization using the geometric face model, we use the following stages [21]:

- Calculating the gray scale of the image I [21]:

$$I= R * 0.299 + G * 0.587 + B * 0.114 \quad (3)$$

Where I express the gray level for a given pixel, R, G and B are the color components of the pixel.

- Calculating the gradient in the horizontal direction

$$\Delta I_y = \partial I / \partial y \, J \quad (4)$$

Where y, J is the index and the horizontal direction, I is the gray scale image.

- Eye axis location (Neyes): The eyes axis location is determined by the maximum of the horizontal projection curve which has a high gradient.

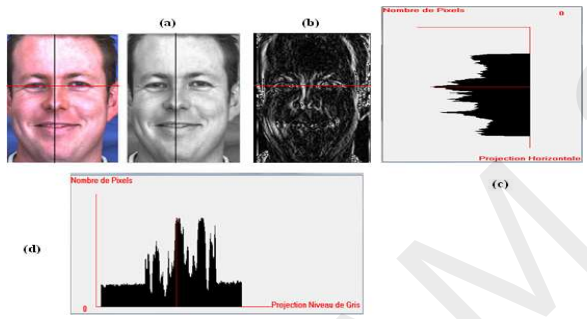
$$HPI_y = \Sigma^{m}_{y=1} \Delta I_y(x,y) \quad (5)$$

$$N_{eyes} = \max (HPI_y) \quad (6)$$

- Median axis location (Ncenter): the median axis location is a vertical line which devises the frontal face in two equal sides. In other words, it is the line passed by the nose. To determine the median axis, we take the position of the highest gray level on the eye axis.

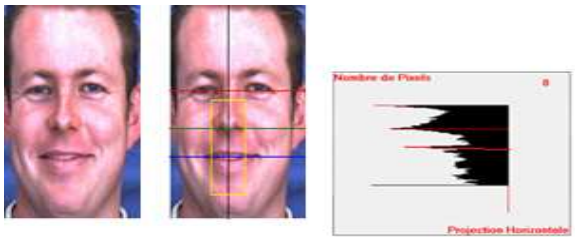
$$N_{center} = \max (I(N_{eyes},y)) \quad (7)$$

Fig. 3. Location of the axe of the eyes and the median



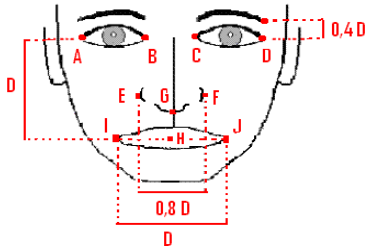
- Nose axis location: search for the highest gradient in a narrow window by horizontal projection, below the eyes axis and around the central axis with a width of pixels Δx (Fig.4) [21].
- Mouth axis location is determined as the same way of nose axis. For the localization of this axis, we look for the maximum value of the projection curve in the low part of the bounding box from nose axis [21].

Fig. 4. Location of the axes of the nostril and mouth



- Application of the geometric model: Once the eyes and mouth axis are located, we use the geometric face model [22].

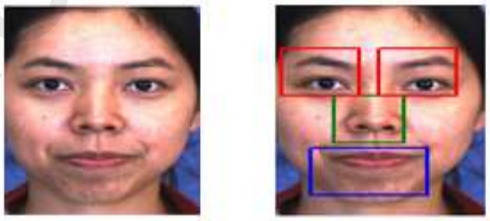
Fig. 5. The geometric model



This model supposes that:

- The distance between the two-eye centers is D.
- The vertical distance between two-eyes and the center of mouth is D.
- The vertical distance between two-eyes and the center of the nostrils is 0.6D.
- The width of the mouth is D.
- The width of nose is 0.8D.
- The vertical distance between eyes and eyebrows is 0.4D.

Fig. 6. Facial features localization



B.2) Facial features point's selection

From the detection windows obtained for the eyes, eyebrows, nose and lips, we use these Region of Interest (RoI) windows to obtain the corner point features, which will be used to fed as feature input vectors to the Neural Network. For this purpose, the Harris corner point detector is implemented. The Harris corner detector's main activity is to evaluate the change in the intensity of individual image windows (selecting a square subset of constant-size pixels) when shifting them in all directions. The algorithm uses the response R of the following local structure matrix:

$$M=W(x,y) \Theta (I^2_x \;\; I_xI_y / I_xI_y \;\; I^2_y) \quad (8)$$

Where: W is a Gaussian window, Θ is the convolution product, Ix and Iy are the first derivatives of the image according to x and y. They are calculated from the convolution of the image with the Sobel filter:

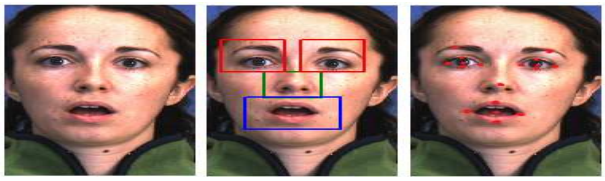
$$I_x = [1\ 2\ 1\ ;\ 0\ 0\ 0\ ;\ 1\ 2\ 1] \quad ; \quad I_y = [\ -1\ -2\ -1;\ 0\ 0\ 0\ ;1\ 2\ 1]$$

The following algorithm shows the steps of the detection of the corners points by Harris method: [23]

- Compute Ix and Iy over the entire image I(x,y)
- For each image point p:
  a) From the matrix M defined by the formula (8);
  b) Compute Det (M), and Trace (M), the determinant and the trace of M;
  c) Compute the response R = Det(M)- k (Trace(M))^2; k ϵ [0.04 – 0.06];
  d) If R> threshold, save the p into a list C;
- Sort C in decreasing order of R;
- Scan the sorted list from top to bottom. For each current point, p, delete all points appearing further in the list which belong to the neighborhood of p.

The Fig.7 shows the feature corner point extraction:

Fig. 7. Feature corner point extraction



The following Fig.8 and Table I present the selected feature corner points

Fig. 8. Facial features points selection

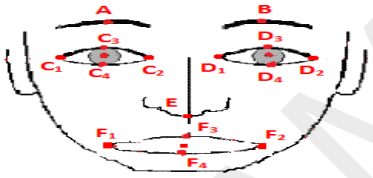


TABLE I. FACIAL FEATURES POINTS SELECTION

Facial Points Selection	Feature point number	Designation
Eyebrow	2 points	A: Left eyebrow center B: Right eyebrow center
Left eye	5 points	C1: Left eye left corner C2: Left eye right corner C3: Left eye upper corner C4: Left eye lower corner C5: Left eye Center
Right eye	5points	D1: Right eye right corner D2: Right eye left corner D3: Right eye upper corner D4 : Right eye lower corner D5: Right eye center
Nose	1 point	E1: Center nose
Mouth	5 points	F1: Lip left corner F2: Lip right corner F3: Lip upper corner F4: Lip lower corner F5: Lip center

C. Finding feature vectors

Now that we have the feature points extracted from the input image, the next step is to decide the input feature vectors that need to be fed into and used to train our neural network. On the basis of these input feature vectors, the Neural Network will hence learn from these and use these inputs to make classification decisions of the final output emotion. The chosen feature vector inputs are as seen in Table II.

TABLE II. INPUT FEATURE VECTORS

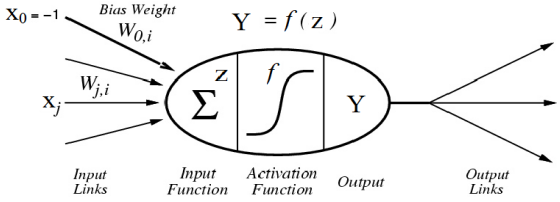
Formula	Definition
Left Eye Height	V1=C1-C2
Left Eye Width	V2 = C4 – C3
Right Eye Height	V3 = D1 – D2
Right Eye Width	V4 = D3- D4
Left eyebrow center to right eyebrow center distance	V5 = A – B
Left eyebrow center to left eye center distance	V6 = A – C5
Right eyebrow center to right eye center distance	V7= B-D5
Lip Height	V8= F1- F2
Lip Width	V9=F3-F4
Left eye right corner to right eye right corner distance	V10= C2-D1
Left eye center to nose center distance	V11= C5-E1
Right eye center to nose center distance	V12=D5-E1
Nose center to lip center distance	V13=E1-F5
Left eye center to lip center distance	V14=C5-F5
Right eye center to lip center distance	V15=D5-F5
Left eye center to right eye center distance	V16=C5-D5

D. Training the neural network

D.1) Neural networks and back propagation

Neural networks have proved their ability in the recent past to deliver simple and powerful solutions in areas relating to signal processing, artificial intelligence and computer vision. A neural network is represented by weighted interconnections between layers of nodes or neurons. Fig.7 shows the structure of an artificial neuron.

Fig. 9. The artificial neuron



Where w_j denotes the weight of the connection linking the neuron j to the input i

The back propagation is a gradient descent technique, which is used to adjust the weights and bias of the network to minimize the quadratic error between the network output and the desired output. At each input / output pair, an error is calculated. Then weights and biases are changed online on the network. These calculations are repeated until the stopping criterion is obtained.

So the back propagation consists in calculating the error of the neurons of a layer from the error of the neurons of the

subsequent layers, the calculation starts at the desired output layer of the neurons:

$$E = \frac{1}{2} \Sigma_k (d_k - y_k^{[s]})^2 \quad (9)$$

Where d_k: is the desired output associated with the input vector x_k,

$y_k^{[s]}$: is the output obtained on the last layer at time t and E : is the cumulated error for k couples presentations (d_k,x_k).

D.2) Neural network configuration

The MLP Neural Network consists of three layers: an input layer, a hidden layer and an output layer.

- The number of neurons in the input layer depends on the number of characteristics used [V1….V16]: 16 neurons.
- The output neurons depend on the number of emotions to be identified: 5 neurons.
- The number of intermediate neurons is initially calculated by the following rule:

NBC = 2*√NBE+NBS= 2*√16+5=10 neurons.

Where NBC: Number of Hidden Neurons, NBE: Number of Entered Neurons and NBS: Number of Output Neurons.

Learning is supervised because the user provides the network with input-output pairs, then he teaches the network all of these pairs by a learning method such as the back propagation of the gradient which tries to reduce the difference between the effective output of the network and the desired output. Learning is complete when all input-output pairs are recognized by the network.

The activation of each neuron is calculated by the sigmoid function given by the following expression:

$$a_i = 1/1+e^{-Neti} \quad (10)$$

$$Net_i = \Sigma^n_{k=0} a_k * w_{ki} \quad (11)$$

Where: Net_i is the weighted activation of neuron i, Wk_i is the weight of the link between neuron i and neuron k, a_i is the output of the network.

In the MLP, the weight assignment is random, and the weights are adapted little by little by the learning process. All training examples are processed until the perceptron classifies them correctly.

For each example, the weights are revised and the correction to be applied to the link weights and neuron thresholds can then be determined using the following rule:

$$W_i = W_i + \Delta W_i \quad (12)$$

$$\Delta W_i = \vartheta (t - O) x_i \quad (13)$$

Where ϑ: the learning step, t: the desired output, O: the output obtained.

After the process of parameter tuning, the neural network was configured by: the optimal learning rate = 0.4, the target of error = 0.001.

IV. Experimental results

To assess the validity and efficiency of our approach, we experimented with the Bi-modal Face and Body Gesture (FABO) Database [1] (Fig.10). The results of the test have been presented as a confusion matrix as shown in Table III, Forty (40) images are trained of distinctive samples and seventy (70) images are tested, the overall efficiency for this condition is 74%.

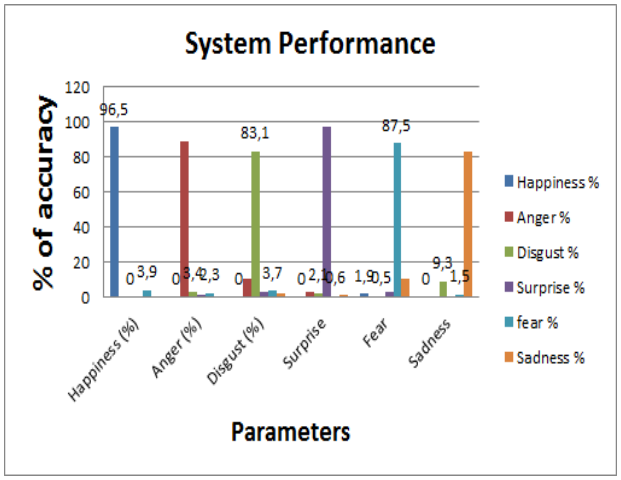
Fig. 10. Sample of images from the FABO facial expression database

TABLE III. Confusion matrix of emotion classification

I/O	Hap[a] (%)	Ang[b] (%)	Dis[c] (%)	Sur[d] (%)	Fea[e] (%)	Sad[f] (%)
Happiness %	96,5	0	0	0	1,9	0
Anger %	0	88,3	10,2	3,2	0,6	0,9
Disgust %	0	3,4	83,1	2,1	0,5	9,3
Surprise %	0	1,6	2,8	97,1	3,1	0,6
fear %	3,9	2,3	3,7	0,6	87,5	1,5
Sadness %	0	0,9	2,4	1,1	10,5	82,4

[a.] Happiness [b.] Anger [c.] disgust [d.] surprise [e.] fear [f.] sadness

Fig. 11. System performance

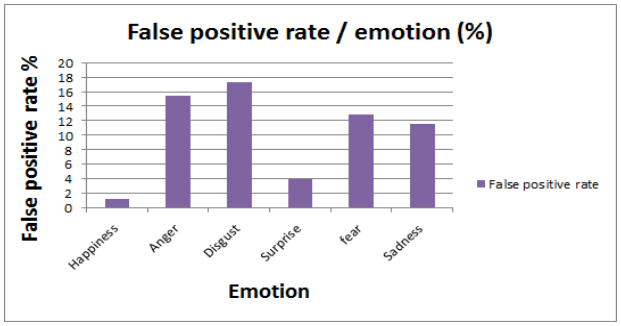


The results of the false positive detection rates per emotion are shown in Table IV.

TABLE IV. FALSE POSITIVE RATE PER EMOTION

Emotion	False positive rate
Happiness	1,2 %
Anger	15,5 %
Disgust	17,3 %
Surprise	3,9 %
fear	12,8 %
Sadness	11,6 %

Fig. 12. False positive rate detection rate per emotion



From Table III and Table IV, we see that the emotions of happiness and surprise are being detected relatively well with 96%, 97% plus true positive success rates, which indicates that the facial reactions when a person is happy and surprised are more uniform than the other emotions leading to high success rates. The emotions anger, sadness, disgust and fear have their success rates in the 83-88% range and have an overlap among almost all of the other emotions indicating that people may have the same facial reaction for two different emotions as well as having different facial reactions for the same emotions, which leads to the overlap among these classified emotions.

V. CONCLUSION

Facial expression recognition has attracted more and more attention duo to its important applications in a wide range of areas.

The application areas of facial expression recognition are increasing and this requires accurate recognition rate on real data. This paper presents a simple approach to automatic facial expression recognition. The proposed system is able to automatically perform human face detection, feature point extraction and facial expression recognition from image sequences. The extraction of facial features sometimes is a very challenging task and it usually takes a lot of computations to extract precise facial features.

Our future aim is the realization and set up of a facial expression recognition system based on a new method to detect the emotional state of the person that measure emotions accurately.

REFERENCES

[1] http://mmv.eecs.qmul.ac.uk/fabo/
[2] A. Mehrabian, "Communication without words," Psychology Today 2(4), 53–56, 1968.
[3] W. Zhao, R. Chellappa, A. Rosenfeld and P.J. Phillips "Face recognition: A literature survey," CVL Technical Report, University of Maryland, October 2000.
[4] P. Viola, M.Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2001.
[5] M. J. Black ,Y. Yacoob, "Recognizing Facial Expressions in Image Sequences Using Local Parametrized Models of Image Motion," International Journal of Computer Vision, Vol. 25, Number 1, pp. 23–48, 1997.
[6] J. Cohn, A. Zlochower, and J. J. Lien and T. Kanade, "Feature-point Tracking by Optical Flow Discriminates Subtle Differences in Facial Expression," Proceedings of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 396–401, 1998.
[7] C. Padgett, G. Cottrell and R. Adolphs, "Categorical Perception in Facial Emotion Classification, Siggraph proceedings," pp. 75–84 1998.
[8] H. Hong, H. Neven and C. von der Malsburg, "Online Facial Expression Recognition based on Personalized Gallery," Intl. Conference on Automatic Face and Gesture Recognition, IEEE Comp. Soc, pp. 354–359, 1998.
[9] T.F. Cootes, G.J. Edwards, C.J. Taylor, "Active Appearance Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 681–685, 2001.
[10] C.L. Huang et Y.M. Huang, "Facial Expression Recognition Using Model-Based Feature Extraction and Action Parameters Classification," Journal of Visual Communication and Image Representation, volume 8 (3), pages 278-290, Septembre 1997.
[11] k. Ghanem, "Reconnaissance des expressions Faciales à base d'informations Vidéo ; Estimation de l'intensité des expressions faciales,"Thèse, Faculté des sciences de l'ingénieur, Algérie, Octobre 2010.
[12] M.J. Lyons , S. Akamatsu, "Coding facial expressions with gabor wavelets," IEEE International Conference on Automatic Face and Gesture Recognition, pages 200–205, Avril 1998.
[13] M.J. Lyons, J. Budynek, S. Akamatsu, "Automatic Classification of Single Facial Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 21 (12), pages 1 357–1 362, Décembre 1999
[14] M.F. Valstar, H. Gunes et M. Pantic, "How to distinguish posed from spontaneous smiles using geometric features," ACM International Conference on Multimodal Interfaces, pages 38-45, 2007
[15] Z. Zhang, M. Lyons, M. Schuster et S. Akamatsu, "Comparison between Geometry-Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron", IEEE International Conference on Automatic Face and Gesture Recognition, pages 454–459, Avril 1998.
[16] I. Cohen, N. Sebe, L. Chen, A. Garg et T.S. Huang, "Facial expression recognition from video sequences : temporal and static modeling," Computer Vision and Image Understanding : Special issue on face recognition, volume 91, pages 160–187, 2003.
[17] F. Davoine, B. Abboud et M. Dang, "Analyse de visages et d'expressions faciales par modèle actif d'apparence," Traitement du Signal, Volume 21 (3), pages 179-193, février 2004.
[18] Z. Hammal, A. Caplier et M. Rombaut, "A Fusion Process Based on Belief Theory Classification of Facial Basic Emotions," International Conference on Information fusion, volume 1, Juillet 2005.
[19] C. Garci, G.Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis," IEEE Transactions on Multimedia, 1(3), p.264-277, September 1999.
[20] M.J. Jones, J. M. Rehg "Statistical Color Models with Application to Skin Detection," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, volume 1, page 1274, Fort Collins, Colorado, June 1999.
[21] F.Aabdat, C.Maaoui et A.Pruski, "Suivi du gradient pour la localisation des caractéristiques faciales dans des images statiques," Colloque GRETSI, Troyes 11-14 septembre 2007.
[22] F.Shih, C. Chuang, "Automatic extraction of head and face boundaries and facial features," Information Sciences 158: 117–130, 2004.
[23] S. El Kaddouhi, A. Saaidi and M. Abarkan, "Eye detection based on the Viola-Jones method and corners points," Multimed Tools Application, DOI 10.1007/s11042-017-4415-5, 2017

Acknowledgment-based Approaches Dealing Against Packet Dropping Attack in MANET: A Survey

Mahdi Bounouni
Faculty of Law and Political Sciences
University of Setif 2, Algeria
LaMOS Research Unit
University of Bejaia, Algeria
Bounouni@gmail.com

Louiza Bouallouche-Medjkoune
LaMOS Research Unit
Faculty of Exact Sciences
University of Bejaia, Algeria
louiza_medjkoune@yahoo.fr

***Abstract*— Mobile ad hoc network (MANET) is a collection of mobile nodes that are self-organizing and able to communicate without relying on any existed infrastructure. In MANET, most of routing protocols rely on the cooperation of all nodes to discover paths and to relay data packets. However, since nodes are not controlled by single authority and are resource-constrained, they may launch packet-dropping attack by behaving maliciously or selfishly. Several reputation approaches have been proposed to cope against packet dropping attack, including promiscuous-based approaches and acknowledgment-based approaches. This paper surveys the current state-of-the-art of the acknowledgment-based approaches, classifies them into two categories according to their ability of detection. Thus, we make a comparative study of these approaches following a set of comparison criteria. The main purpose is to reveal the strengths and weaknesses of each approach in order to direct the readers of this paper towards the future trends of this category of reputation approaches.**

Keywords— MANET, security, packet dropping attack, malicious, selfish, acknowledgment

I. INTRODUCTION (*HEADING 1*)

An ad hoc Mobile Network (MANET) is a wireless network that consists on a set of mobile nodes able to communicate without relying on an existing infrastructure or centralized administration. In MANET, the network tasks are performed by nodes without relying on any existing infrastructure based on the assumption that all nodes are willing to cooperate [1]. Therefore, the well-functioning of these tasks depends on the cooperation of all nodes. However, this cooperation cannot be ensured due to the specific characteristics of MANET, especially the constraint of limited-resource of nodes and the lack of central authority to manage nodes.

In MANET routing protocols, each node plays the role of a router and a host. To be able to communicate with a distant node, a node relies on its neighbors to forward its packets based on the assumption that all of its neighbors behave correctly and in accordance with the routing rules. However, a node can deviate its behavior by dropping packets destined to be forwarded, either to disrupt forwarding activities (malicious behavior), or to preserve its resources (selfish behavior). These behaviors are known as the packet dropping attack. In this attack, nodes behave inconsistently with the routing rules. Without appropriate countermeasures, packet dropping attack can cause a significant degradation of network performance.

Several surveys [2-4] have been proposed to address the security issues of routing protocols in MANET. Almost of them study and discuss various approaches proposed to cope against different types of attacks. Then, it is interesting to focus on a specific attack and a particular category of approaches. In this paper, we focus on the acknowledgment-based approaches proposed to thwart the packet dropping attack. Thus, we make comparison study between these approaches in order to highlight their effectiveness and weaknesses which constitutes future trends of the audience of the paper. This paper is a scoping review that focuses on a specific attack, while at the same time addresses a specific category of approaches with further details. The aim is to facilitate for researchers to grasp acknowledgment-based approaches.

The remainder of the paper is structured as follows. Section 2 discusses packet dropping attack. The background information required for understanding this paper is given in section 3. In section 4, we review the acknowledgment-based approaches. Section 5 discusses the acknowledgment approaches by performing a comparative study. Section 6 concludes the paper.

II. PACKET DROPPING ATTACK: AN OVERVIEW

The packet dropping attack consists on dropping packets destined to be routed, including the routing packets and data packets. Depending on the attacker's behavior, we classify the packet dropping attack into two categories: the packet dropping attack in which the attacker behaves maliciously, and the packet dropping attack in which the attacker behaves selfishly. In this paper, we use the term uncooperative node to refer to both selfish and malicious nodes.

A. Malicious behaviors

A malicious attacker is intended to drops the packets in order to disrupt the process of forwarding packets and to cause damage. It can act with different strategies to achieve its goal. We present two strategies of malicious nodes related to the DSR protocol [5]:

1) Strategy (1): an attacker can participate cooperatively in the route discovery process. It routes correctly all the RREQs and RREPs packets passing through it without any alteration or modification. But, once the attacker is implied in a forwarding route, it drops all data packets passing through it.

2) Strategy (2): this strategy is similar to strategy (1), except that instead of dropping all the data packets passing through it, the attacker drops only a fraction of them.

B. Selfish behaviors

Selfish behavior is different from malicious behavior. A selfish node is not intended to attack other nodes of the network. It uses network resources to route its packets, but it refuses to route packets for the benefits of other nodes in order to preserve its resources (since forwarding packets consumes resources). The authors in [6] have defined three types of selfish nodes related to the DSR protocol [5]:

1) Selfish type (1): Selfish nodes type (1) participate in route discovery and route maintenance processes. But, once they are involved in a forwarding route, they drop all the data packets passing through them.

2) Selfish type (2): Selfish nodes type (2) refuse to cooperate in the route discovery process in order to avoid being involved in the discovered routes. They drop all RREQ packets to avoid being included in any forwarding route. They use their resources only for their own transmissions.

3) Selfish type (3) :Selfish nodes type (3) behave according to their residual energy levels. They behave in a cooperative manner by forwarding all the packets destined to be routed, if their energy levels are higher than a certain threshold T1. However, if their energy levels are between both thresholds T and T2 (with T1 > T2), they behave like nodes of type (1). But, if their energy levels are below the T2 threshold, they behave like the selfish type (2).

III. SURVEY BACKGROUND

In the literature, several approaches [7-9] have been proposed to secure mobile ad hoc network. Most of these approaches rely on the use of cryptographic technique such as digital signature, symmetric and asymmetric encryption. Cryptographic approaches aim to ensure some security properties such as confidentiality, data integrity, entity authentication [10]. Although these approaches can thwart several types of attacks such as Fabrication and the falsification of routing packets, impersonation [10], they fail to deal against packet dropping attack.

Reputation approaches have been proposed to detect and punish nodes dropping packets. They are based on monitoring the forwarding activities of neighbor nodes. Each node monitors its neighboring nodes and it calculates their reputation values according to their behaviors. The reputation value of a node increases if it forwards correctly a data packet. Otherwise, it decreases if it drops a data packet. Note that the reputation value of a node reflects its trustworthiness. If the reputation value of a node falls below a certain threshold, the node is considered uncooperative. Based on the monitoring technique used, we can classify reputation approaches into two categories: promiscuous-based approaches and acknowledgment-based approaches.

The promiscuous-based [11] approaches rely on the use of the promiscuous mode to monitor the behavior of neighbor nodes. Using this mode, if a node A is in the transmission range of node B, the node A can overhear the communications of the node B even in the case when these communications do not involve this node directly. One of the most reputable approach of these category is the watchdog approach [11]. The watchdog monitors and checks whether the next node on the forwarding route forwards correctly the data packet recently sent. Many security approaches [11-14] employ the watchdog approach to unmask nodes dropping packets. However, although the use of the watchdog approach allows to detect uncooperative nodes, it suffers from several limitations [11] and All security approaches using this technique inherit these limitations. The approaches using this technique fail to detect malicious nodes in the presence of: ambiguous collision, receiver collision, limited transmission power, false accusation attack, collusion attack and partial dropping attack.

To overcome the limitations of promiscuous-based approaches, the acknowledgment-based approaches [15-20] have been proposed. These approaches permit to extend the range of neighborhood monitoring to two hops, compared to promiscuous-based approaches that enable to monitor only one-hop neighbors. We discuss this category of approaches with further details in the next section.

IV. ACKNOWLEDGMENT-BASED APPROACHES

As described in the previous section, the acknowledgment-based approaches have been proposed to deal against packet dropping attack, while at the same time overcome the weaknesses of promiscuous-based approaches. In the literature, several acknowledgment-based approaches have been proposed. We can classify the acknowledgment approaches into two categories: Link-detection approaches and node-detection approaches. This classification is made based on the fact that some approaches enables nodes to detect only uncooperative link and the others approaches can detect uncooperative nodes.

A. *Link-detection approaches*

One of the most reputable approaches of this category is TWOACK approach [15]. In this approach, the authors have introduced a new type of packet called TWOACK. This latter is used to ensure that the node located at two hops has correctly received the data packet. For each data packet received, a node sends a TWOACK packet to the node located at two hops in the opposite direction of the forwarding route. We take the triplet of nodes $ A, B, C $ as an example to illustrate the functioning of this approach (see Fig. 1). Then, for each data packet received, node C sends back a TWOACK packet to a node A. Each node maintains a counter per transmission link (in our example, node A maintains a counter for link B-C). This counter represents the number of unacknowledged data packets with TWOACK packets. If a data packet is unacknowledged before the expiration of the timer T, the counter is incremented. If the counter exceeds the tolerated threshold, the monitored link is declared as uncooperative.

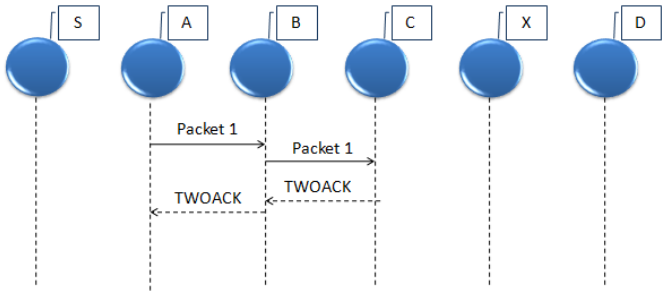


Figure 1. Functioning of TWOACK approach

The authors in [15] proposed the S-TWOACK approach (Selective-TWOACK) in order to reduce the overhead due to the exchange of TWOACK packet. In the S-TWOACK approach, instead of acknowledging all data packets, a node sends back a single TWOACK packet for multiple data packets. Although this approach can reduce overhead compared to the TWOACK approach, it introduces additional detection time because it takes much times to detect uncooperative links.

The 2ACK approach has been proposed in [16]. The functioning of the 2ACK approach is similar to the TWOACK approach [15]. The main difference between these two approaches is that the 2ACK approach acknowledges only a fraction of the data packets and it can prevent nodes from forging 2ACK packets by employing the digital signature technique.

Based on the TWOACK approach, the AACK [17] approach has been proposed to secure multimedia traffic in an ad hoc mobile network. The purpose of this approach is to reduce the overhead due to the exchange of TWOACK packets, especially when there is no uncooperative activities along a forwarding route. The AACK approach is the result of a combination of two modes: AACK and TACK. The AACK mode is similar to the traditional acknowledgment approach in which the destination sends back an ACK packet to the source for each data packet received. In AACK mode, if the source does not receive an ACK for a data packet after a predefined time period, the source switches to TACK mode. The functioning of the TACK mode is similar to the TWOACK approach. The TACK mode is used until the destination receives a data packet correctly. At start-up, the AACK mode is used by default.

The EAACK approach [18] is an extension of the AACK approach that aims to prevent nodes from forging TACK packets and to thwart the false accusation attack. It is the result of combination of three modes: ACK, S-ACK, MRA. Both ACK and S-ACK modes are identical to the two AACK and TACK modes introduced in the AACK approach, except the use of both RSA and DSA digital signature. In the EAACK approach, all acknowledgment packets should be digitally signed before they are sent out and they should be verified. Thus, the authors introduced MRA (Misbehavior Report Authentication) mode to cope against malicious nodes that wrongly accuse cooperative nodes to be malicious. To initiate the MRA mode, the source node first searches an alternative route to the destination node. Then, to check whether the report received is valid, the source node sends MRA packet to the destination using the alternative route. The purpose is to check the safety of the misbehavior report. Then, if the packets have been received, the source concludes that it is a false report. Otherwise, the report is accepted and the reported link is marked as malicious.

Although acknowledgment-based approaches can overcome some limitations of promiscuous-based approaches, they also suffer several weaknesses. This categoy of approaches can only identify uncooperative links instead of uncooperative nodes which give for them more chance to drop a lot of data packets. Uncooperative nodes can exploit the features of MANET (especially the mobility) to involve themselves in multiple forwarding paths in order to damage the forwarding activities of data packets. Thus, these approaches assume that all nodes are willing to cooperate in the route discovery process. Then, the nodes dropping REEQ and RREP packets are never monitored and detected and their packets are always relayed by cooperative nodes.

B. *Node-detection approaches*

The authors in [1] proposed the NHACK approach (New Hybrid Acknowledgment). The aim is to identify and punish uncooperative nodes dropping data packets, while at the same time to stimulate selfish nodes to route RREQ and RREP packets. Based acknowledgment monitoring technique [15], a new reputation calculation method is incorporated. This method permits to quantify the behavior of a node with a single reputation value instead of quantifying the reputation values of forwarding links. The computed reputation value is used to penalize nodes having reputation values below the tolerated threshold. Thus, the reputation values of nodes are incorporated in the route discovery process in order to enable nodes to select reliable forwarding paths that involve only high-reputed nodes. Thus, to stimulate selfish nodes to cooperate in the route discovery process, the contribution each node provided in the route discovery process is used as an incentive of cooperation. The contribution reflects the number of RREQ and RREP packets routed by a node within its neighborhood. Although the NHACK approach introduces a reputation constraint in the route discovery process, low-reputed can be implied in forwarding paths because the authors take into account the

reputation values of the routes instead of the reputation value of each node involved in a forwarding path.

In [19], HAPS (Hybrid Acknowledgement Punishment and Stimulation) approach have been proposed. In this approach, the authors aim to Enhance the performance of EAACK approach by punishing uncooperative nodes more severely, ensure a fairness between nodes and to thwart false dissemination attack. HAPS approach is structured around three modules: monitor, reputation, exclusion. The monitor component monitors the behaviors of neighbors nodes by monitoring their forwarding links using the acknowledgment monitoring technique [18]. The reputation module combines the direct and indirect reputation of the nodes to come up a single reputation value of a node. The direct reputation of a node reflects its trustworthiness in all forwarding links in which is involved. It is computed by aggregating the recommendation of the monitor module using the Dempster Shafer theory [20]. Thus, HAPS approach enables nodes to exchange their recommendations about their neighbours in order to computer their indirect reputation values. The exclusion module punishes nodes having reputation values smaller than the reputation threshold. The main drawback of this technique is how to ensure that nodes forward correctly routing packets? And, how to motivate these nodes to cooperate, and to prevent them using the network resources if they refuse?

In [21], the authors proposed an acknowledgment-based punishment and stimulation scheme, called APS. APS approach aims to punish uncooperative nodes dropping data and routing packets while at the same time motivate selfish nodes to cooperate using credits. The APS approach is structured around four interactive modules. The monitoring agent is responsible for monitoring the correct forwarding of routing and data packets. The reputation module calculates the direct and indirect reputation of each neighbor. The direct reputation value of a node is the result of the combination of data and routing reputation. The APS approach allows nodes to share periodically their recommendations about other nodes. And, the combination of all recommendations is done using the CIF (Combined trust on Importance Factor) rule [22]. The stimulator module motivates nodes to cooperate using credits. In this approach, nodes are stimulated to fully cooperate in order to be able to send their packets freely or with a low price. The isolation module punishes malicious nodes and selfish nodes. Although this approach permits to detect uncooperative nodes and to motivate selfish nodes, they are unable to detect two colluding uncooperative nodes along a forwarding route that try to hide their misbehavior by generating fake acknowledgment packets.

V. DISCUSSION

In this section, we make a comparative study of all acknowledgment-based approaches presented in the previous section.

A. *Addressed Limitations*

The acknowledgment-based approaches have been proposed to overcome the limitations of promiscuous-based approaches namely:

- ✓ Ambiguous Collision: This limitation prevents the node from overhearing the forwarding activities of the next hop node on a forwarding route;
- ✓ Receiver collision: in this case, the node can overhear the communication of the of the next hop node, but it is not able to confirm the good reception of the packet;
- ✓ Limited transmission power: a node can adjust its transmission power so that the signal is loud enough to be overheard by the previous node, but it is too weak to be received by the next node.
- ✓ False accusation attack: a uncooperative node may falsely accuse other cooperative nodes to be malicious in order to evict to forward packets.
- ✓ Collusion attack: When two uncooperatives nodes succeed in a forwarding route, they can collaborate in order to hide their dropping activities (misbehavior's);
- ✓ Partial dropping attack: An uncooperative node can circumvent the watchdog by dropping only a fraction of packets.

A summary of the limitations resolved by each approach is presented in Table. 1.

B. *Approaches evaluation challenges*

In order to compare different acknowledgment approaches, a summary of the characteristics of the surveyed approaches is described in Table. 2. We highlight the strong and the weak points of each approach. The features of each approach are highlighted using on the following criteria:

- ✓ Selfish detection: Selfish node type to which the approach can deal with.
- ✓ Malicious detection: Malicious strategies to which the approach can cope with.
- ✓ Sharing recommendation: is the approach enable nodes to share their recommendations about their neighbors in order to compute their indirect reputation values?
- ✓ Redemption: check if the approach allows to reintegrate the nodes considered as malicious in the network (a second chance).
- ✓ Reaction: Once a node is considered malicious, the reaction mechanism indicates whether the approach punishes them by refusing to forward their packets.

TABLE I. RESOLVED LIMITATIONS

Approaches	**Ambiguous collision**	**Receiver collision**	**Limited transmission power**	**False accusation attack**	**Collusion attack**	**Partial dropping attack**
TWOACK [15]	√	√	√			
S-TWOACK [15]	√	√	√			
2ACK [16]	√	√	√			
AACK [17]	√	√	√			
EAACK [18]	√	√	√			
NHACK [19]	√	√	√			√
HAPS [20]	√	√	√	√		√
APS [21]	√	√	√	√		√

TABLE II. EVALUATION OF ACKNOWLDGMENT-BASED APPROACHES

Approaches	**Malicious detection**	**Selfish detection**	**Sharing recommendations**	**Redemption**	**Reaction**	**Communication overhead**
TWOACK [15]	(1)	(1)	No	No	No	High
S-TWOACK [15]	(1)	(1)	No	No	No	Low
2ACK [16]	(1) & (2)	(1)	No	No	No	Medium
AACK [17]	(1)	(1)	No	No	No	Medium
EAACK [18]	(1)	(1)	No	No	No	High
NHACK [1]	(1) & (2)	(1) & (2)	No	Yes	Yes	High
HAPS [19]	(1)	(1)	Yes	No	Yes	Medium
APS [20]	(1) & (2)	(1) & (2)	Yes	YEs	Yes	Medium

- ✓ Communication overhead: to detect nodes dropping packets, the acknowledgment approaches rely on the exchange of acknowledgment packets which generates an additional overhead.

VI. Conclusion

In this paper, we have surveyed the state-of-the-art of acknowledgment-based approaches proposed to cope against nodes misbehavior in MANET. The acknowledgment-based approaches have been proposed to address the limitation of the promiscuous-based approaches. We categorized them into two categories according to their detections abilities. There are approaches that enable nodes to detection the malicious links. Other approaches are designed to detect and avoid malicious nodes instead of malicious links. A comparative study between these approaches was conducted to highlight their effectiveness and weaknesses which give some future trends for the reader and audience of the paper.

As future work, we plan to study and discuss the stimulation approaches proposed to stimulate the nodes cooperation. Thus, we intend to propose an approach to thwart the collusion attack that all of acknowledgment approaches fail to detect. Thus, one of major limitation of these approaches is their high communication overhead. Therefore, proposing a technique to minimize the communication overhead of these approaches without affecting their performance constitute a challenge.

VII. References

[1] M. Bounouni and L. Bouallouche-Medjkoune, "A Hybrid Stimulation Approach for Coping Against the Malevolence and Selfishness in Mobile Ad hoc Network," Wireless Personal Communications, vol. 88, no. 2, pp. 255–281, 2015.

[2] R. H. Jhaveri, S. J. Patel, and D. C. Jinwala, "DoS Attacks in Mobile Ad Hoc Networks: A Survey," Second International Conference on Advanced Computing & Communication Technologies, 2012.

[3] K. Govindan and P. Mohapatra, "Trust Computations and Trust Dynamics in Mobile Adhoc Networks: A Survey," IEEE Communications Surveys & Tutorials, vol. 14, no. 2, pp. 279–298, 2012.

[4] Liu, G., Yan, Z., Pedrycz, W.: Data collection for attack detection and security measurement in mobile ad hoc networks: A survey. Journal of Network and Computer Applications, (2018).

[5] D. B. Johnson and D. A. Maltz, "Dynamic Source Routing in Ad Hoc Wireless Networks," The Kluwer International Series in Engineering and Computer Science Mobile Computing, pp. 153–181.

[6] P. Michiardi, and R. Molva, "Simulation-based analysis of security exposures in mobile ad hoc networks" European Wireless Conference, pp. 15-17, 2002

[7] M. G. Zapata and N. Asokan, "Securing ad hoc routing protocols," Proceedings of the ACM workshop on Wireless security - WiSE 02, 2002.

[8] P. Papadimitratos and Z. J. Haas, "Secure message transmission in mobile ad hoc networks," Ad Hoc Networks, vol. 1, no. 1, pp. 193–209, 2003.

[9] K. Vijayalakshmi and S. Nagamalai, "Self-healing key distribution for self-organised group management in MANET," International Journal of Mobile Network Design and Innovation, vol. 6, no. 4, p. 212, 2016.

[10] D. Djenouri, L. Khelladi, and A. Badache, "A survey of security issues in mobile ad hoc and sensor networks," IEEE Communications Surveys & Tutorials, vol. 7, no. 4, pp. 2–28, 2005.

[11] S. Marti, T. J. Giuli, K. Lai, and M. Baker, "Mitigating routing misbehavior in mobile ad hoc networks," Proceedings of the 6th annual international conference on Mobile computing and networking - MobiCom 00, 2000.

[12] S. Bansal, M. Baker, "Observation-based cooperation enforcement in ad hoc networks," arXiv preprint cs/0307012, 2003

[13] S. Buchegger and J.-Y. L. Boudec, "Performance analysis of the CONFIDANT protocol," Proceedings of the 3rd ACM international symposium on Mobile ad hoc networking & computing - MobiHoc 02, 2002.

[14] P. Michiardi and R. Molva, "Core: A Collaborative Reputation Mechanism to Enforce Node Cooperation in Mobile Ad Hoc Networks," Advanced Communications and Multimedia Security IFIP Advances in Information and Communication Technology, pp. 107–121, 2002.

[15] K. Balakrishnan, J. Deng, and P. Varshney, "TWOAK: preventing selfishness in mobile ad hoc networks," IEEE Wireless Communications and Networking Conference, 2005.

[16] K. Liu, J. Deng, P. K. Varshney, and K. Balakrishnan, "An Acknowledgment-Based Approach for the Detection of Routing Misbehavior in MANETs," IEEE Transactions on Mobile Computing, vol. 6, no. 5, pp. 536–550, 2007.

[17] T. Sheltami, A. Al-Roubaiey, E. Shakshuki, and A. Mahmoud, "Video transmission enhancement in presence of misbehaving nodes in MANETs," Multimedia Systems, vol. 15, no. 5, pp. 273–282, 2009.

[18] E. M. Shakshuki, N. Kang, and T. R. Sheltami, "EAACK—A Secure Intrusion-Detection System for MANETs," IEEE Transactions on Industrial Electronics, vol. 60, no. 3, pp. 1089–1098, 2013.

[19] M. Bounouni and L. Bouallouche-Medjkoune, "Hybrid Acknowledgment Punishment Scheme Based on Dempster-Shafer Theory for MANET," Computational Intelligence and Its Applications IFIP Advances in Information and Communication Technology, pp. 436–447, 2018.

[20] T. Chen and V. Venkataramanan, "Dempster-Shafer Theory for Intrusion Detection in Ad Hoc Networks," IEEE Internet Computing, vol. 9, no. 6, pp. 35–41, 2005.

[21] M. Bounouni and L. Bouallouche-Medjkoune, "Acknowledgment-based punishment and stimulation scheme for mobile ad hoc network," The Journal of Supercomputing, vol. 74, no. 10, pp. 5373–5398, 2018.

[22] A. Jesudoss, S. K. Raja, and A. Sulaiman, "Stimulating truth-telling and cooperation among nodes in VANETs through payment and punishment scheme," Ad Hoc Networks, vol. 24, pp. 250–263, 2015.

Mapping and scheduling techniques in NoC: A survey of the state of the art

Djalila Belkebir
Kasdi Merbah Ouargla University, Algeria
belkebir.djalila@gmail.com

Adel Zga
Kasdi Merbah Ouargla University, Algeria
zgaadell@gmail.com

***Abstract*—Mapping and Scheduling steps follow NoC specialization and their role is to implement the given application in to the selected architecture which mainly means to assign and order the tasks and communications of the application in to the resources of the architecture such that the design goals to be optimized. The aim of this paper is to present a detailed survey of the work done in last two decade in the domain of mapping and scheduling of applications on to NoC architectures, also a classification of design-time mapping and run-time mapping on to NoC were presented.**

Keywords—Mapping, Scheduling, Network on Chip, Run-time, Design-time.

I. INTRODUCTION

Due to increasing systems complexity and design productivity gap, the design of future multiprocessor systems-on-chip (MPSoC) slow down [1]. From that, Network on Chip (NoC) is a promising interconnect solution which is a new design paradigm and on-chip architecture which aims to overcome the design problems and performance limitations of current bus-based Systems on Chip (SoC) methodologies [2]. NoC architecture consists of many heterogeneous intellectual property (IP) cores. Where the wires are replaced by a network of shared links and multiple routers exchanging data packets simultaneously.

The last phase is currently assuming more and more interest in the scientific community [3],[4]. Actually, it has a strong impact on typical performance indexes to be optimized. The mapping problem is an instance of constrained quadratic assignment problem which is known to be NP-hard [5]. The search space of the problem increases factorially with the system size. It is therefore of strategic importance to define methods to search for a mapping that will optimize the desired performance indexes (performance, power consumption, quality of service, etc.) with a good tradeoff between accuracy and efficiency. The search for an optimal mapping (henceforward referred to as exploration) is also complicated when the concept of optimality is not limited to a single performance index (or objective) but comprises several contrasting indexes.

The aim of this paper is to survey the most knowing mapping and scheduling technique in design time and run-time, that is organized as follows: in Section 2, key factors on task mapping are presented followed by Section 3 where the related work of mapping and scheduling techniques in NoC are discussed and Section 4 concludes this paper.

II. KEY FACTORS ON TASK MAPPING

Mapping of the IP core onto NoC tile could be done by either assigning a single core onto each tile or by assigning multiple IP core on each of the tile (Scheduling). Each of the mapping procedure has its own merits and demerits. Many factors should be considered during the mapping of task onto NoC tiles, such as:

A. Target architecture

The target architecture is related to whether nodes in the NoC system are heterogeneous or homogeneous. Heterogeneity is the most common case, because this factor may improve system performance in the presence of different kinds of applications. Heterogeneity refers to having several kinds of nodes in the system (i.e., nodes may be different among them).

B. Abstraction level of the application specification

The abstraction level in which applications are described is a key factor in mapping tasks of such applications to the available resources. The first possible approach to this subject is to use Register Transfer Level (RTL). RTL is a valuable tool for modeling and designing complex systems, and often relies on hardware description languages, such as VHDL (VHSIC Hardware Description Language) and Verilog. Such tools allow modeling a part of the NoC system such as the communication system [6]. The second reported approach is based on transaction-level modeling or TLM. Transactions are defined as the event of synchronization or data exchange among system modules. This approach is appealing because it allows performing a functional verification of the system, and the modeling is based on languages such as SystemC. TLM has been used successfully for synthesizing high speed MPSoC systems and for modeling the communications infrastructure of a NoC [6].

C. Figures of merit

This factor refers to the optimization criteria which must be considered along the optimization process related to the mapping stage. Such optimization can be viewed as a solutions space exploration, where each solution represents a single design choice with different values of the objective functions [6]. The task mapping process must find an acceptable solution within the space with allowable and optimized values for such functions. Among the most common figures of merit used for such optimization process, we may find: power consumption, delay

time, mapping time, temperature, mean number of hops across the network, network contention, mean channel occupancy, bandwidth, and so on.

D. Common domain semantic

This is a medium level representation which combines information both from the high-level application description and from the implementation platform. Among the plethora of representations available for these purposes, graph-based approaches are the most common, with instances such as task graphs (TG), communication task graphs (CTG), communication weight graphs (CWG), communication resources graph (CRG), annotated task graphs (ATG), synchronous and asynchronous data flow graphs (SDFG and ADFG), and so on. Some other kinds of such medium– level representations are the Petri Networks (PN), and the Kahn Process Networks (KPN) [6].

E. Topologies

Topology refers to the way in which system nodes are physically interconnected. Topologies may be classified as either regular or irregular. Some instances of common topologies are meshes, torus, rings, and Spidergon ones. Regular topologies are more constrained with respect to the distribution of the connections, which are generated by means of mathematical functions. Irregular Topologies are often the mixture of two or more regular forms, which leads to hybrid, hierarchical or totally irregular topologies.

F. Optimization algorithms

As already mentioned, the mapping stage relies on an optimization process, which searches along the solutions space, the design with a better tradeoff among the chosen figures of merit. The kind of optimization algorithm used for task mapping has a direct impact on the communications nature. For instance, off–line (static) optimization forces to having predictable communication assessments, whilst dynamic algorithms allow a more flexible communication scheme.

III. RELATED WORK OF MAPPING AND SCHEDULING TECHNIQUES IN NOC

There could be a number of taxonomies to classify the mapping methodologies, like target architecture based, optimization criteria based, workload based, etc. Broadly, the methodologies can be classified based on workload scenarios and other taxonomies can be included at some hierarchy in the classification. For static and dynamic workload scenarios, the mapping methodologies perform optimization at design-time and run-time respectively, which has led them to classify as design-time and run-time methodologies respectively. The methodologies target either homogeneous or heterogeneous multi-core systems.

A. Design-time mapping

Design-time mapping methodologies have a global view of the system which facilitates in making better decision about using the system resources. As optimization is performed at design-time, the methodologies can use more thorough system information to make decisions. Thus, a better quality of mapping may be achieved as compared to the runtime mapping methodologies that are normally restricted to a local view where only the neighborhood of the task mapping is considered. Design-time strategies are suited only for mapping predefined set of applications and thus cannot predict dynamic behavior incurred due to the target applications and state of the platform at run-time. This dynamism demands run-time mapping of application tasks to maintain a critical balance between performance and resource optimization. Any disturbance may either lead to digression from expected performance or complete drain of valuable resources. So, it becomes mandatory to devise algorithms which can intelligently distribute the application tasks among processors taking communication overhead, computation load and resource utilization in consideration.

In [7], the author proposed an evolutionary algorithm-based technique. To hide memory latency, prefetching is aggressively performed in the proposed technique. The experimental results show that the overlay overhead significantly was reduced compared to a non-optimized approach. Another work applied a round-robin scheduling techniques to achieve high communication resource utilization [8], the author proposed a switch which employs the latency insensitive concepts. Based on the assumptions of the 2D-mesh network topology constructed by the switch. The algorithm is based on the communication-driven task binding in such a way that the overall system throughput is maximized. The experimental results demonstrate that the overall improvement of the system throughput is 20% during 844 test cases compared to the task binding without considering the communication and contention effect. The author in [9], proposed an algorithm which automatically maps NoC a given set of IP onto a generic regular architecture and constructs a deadlock-free deterministic routing function such that the total communication energy is minimized based on two steps. The first one is the problem formulation of energy- and performance-aware mapping in a topological. An efficient branch-and-bound (BB) algorithm is then proposed to solve this problem. Simulation experiences for a complex video/audio application shows that the average communication energy savings is about 51.7%, compared to an ad hoc implementation.

In [10], the author proposed a complete allocation and scheduling framework. The optimizer implements an efficient and exact approach to allocation and scheduling based on problem decomposition. The allocation subproblem is solved through integer programming while the scheduling one through constraint programming. The two solvers can interact by means of no-good generation, thus building an iterative procedure which has been proven to converge to the optimal solution. Another work in [11], a decomposition-based approach to speed up constraint optimization have been proposed in order to optimize for the schedule length or make-span. In [12], the author proposed a mapping framework called Distributed Operation Layer (DOL), which optimizes for computation and communication time. They integrate an analytic performance analysis strategy into DOL to alleviate the modeling and analysis of systems.

In [13], a new parameter selection scheme for simulated annealing is proposed that sets task mapping specific optimization parameters automatically. The scheme bounds

optimization iterations to a reasonable limit and defines an annealing schedule that scales up with the application and architecture complexity. The presented parameter selection scheme compared to extensive optimization achieves 90% goodness in results with only 5% optimization time, which helps large-scale architecture exploration where optimization time is important. The optimization procedure is analyzed with simulated annealing, group migration and random mapping algorithms using test graphs from the Standard Task Graph Set. Simulated annealing is found better than other algorithms in terms of both optimization time and the result. Simultaneous time and memory optimization method with simulated annealing is shown to speed up execution by 63% without memory buffer size increase. As a comparison, optimizing only the execution time yields 112% speedup, but also increases memory buffers by 49%.

In [14], the author proposed compiler approach that has four major steps: task scheduling, processor mapping, data mapping, and packet routing. In the first step, the application code is parallelized and the resulting parallel threads are assigned to virtual processors. The second step implements a virtual processor-to-physical processor mapping. The goal of this mapping is to ensure that the threads that are expected to communicate frequently with each other are assigned to neighboring processors as much as possible. In the third step, data elements are mapped to the memories attached to CMP nodes. The main objective of this mapping is to place a given data item into a node which is closer to the nodes that access it. The last step of this approach determines the paths between memories and processors for data to travel in an energy efficient manner. The experimental analysis shows that the proposed framework reduces energy consumption about 27.41% on average over a pure performance-oriented application mapping strategy. In [15], the author proposed a methodology consisting of Integer Linear Programming (ILP) formulation to explore efficient mappings. The searches-based approaches provide efficient mapping solutions, but they have high computational costs for large scale problems such as applications with large number of tasks. Different pruning strategies have been incorporated to prune the search space, thereby reducing the computational costs. In [16], an approach was developed that effectively and efficiently allocates execution and storage slack to jointly optimize system lifetime and cost. While exploring less than 1.4% of the slack allocation design space, its approach consistently outperforms alternative slack allocation techniques to find sets of designs within 1.4% of the lifetime-cost Pareto-optimal front.

In [17], the author proposed an algorithm to solve the allocation and scheduling problem based on constraint programming. He introduced a number of search acceleration techniques that significantly reduce run-time by pruning the search space without compromising optimality. The solver has been tested on a number of non-trivial instances and demonstrated promising run-times on SDFGs of practical size and one order of magnitude speed-up the fastest known complete approach. In [18], the author proposed a thermal-aware system analysis method that produces mappings with a lower peak temperature of the system, leading to reliable system design. And in [19], the author considered the number of software pipeline stages to map streaming applications on SPM-based embedded multi-core system. The proposed method scales well over a wide range of cores and SPMs.

In [20], a unified task scheduling and core mapping algorithm called UNISM have been proposed for different NOC architectures including regular mesh, irregular mesh and custom networks. First, a unified model combining scheduling and mapping is introduced using MILP. Then, a novel graph model is proposed to consider the network irregularity and estimate communication energy and latency, since the number of network hops is not accurate enough for irregular mesh and custom networks. To make the MILP-based UNISM scalable, a heuristic is employed to speed up the method. Experimental results show that more than 15% and 11.5% improvement on the execution time is achieved with similar energy consumption on average for regular mesh NOC compared to other works. For irregular and custom NOC, the improvement is 27.3% and 14.5% on the execution time with 24.3% and 18.5% lower energy. In [21], The mapping strategy is represented by chromosomes, that affect tasks with hard real-time requirements to cores in order to minimize power dissipation under timing constraints. The optimization is multi-objective that combines power estimation macro-models and real-time schedulability analysis.

In [22], the author presented a constructive heuristic to statically map applications on two-dimensional mesh-connected NoC. The approach corresponds to a design time decision of attachment of cores to the routers. The mapping results, in terms of overall communication cost metric, have been compared with many well-known techniques reported in the literature and also with an exact method built around ILP. In [23], the author addressed the problem of task mapping in the context of multiprocessor applications with stochastic execution times and in the presence of constraints on the percentage of missed deadlines.

In [24], BB-based exact mapping (BEMAP) algorithm, for mapping real-time embedded applications on the NoC architecture have been proposed. The BEMAP optimizes the latency and throughput of the NoC system and minimizes power consumption under the bandwidth constraint. This method utilizes the modular exact and systematic search optimization techniques to obtain a multi-objective optimized solution to the mapping problem of the NoC designs. The proposed algorithm exploits the state-of-the-art BB algorithm, in order to obtain optimized results against its competitors. Simulation shows that the proposed algorithm achieves up to 19.93% savings in power consumption and 61.10% improvement in network latency for two-dimension mesh and torus topologies.

In [25], the author proposed a strategy to increase the thermal safety of NoC-based systems by a graceful decrease in communication cost and ILP formulation to deal with the problem. To overcome huge computational overhead of ILP, another solution strategy, based on meta-heuristic technique, Particle Swarm Optimization (PSO) is also proposed. Several innovative augmentations have been introduced into the basic PSO to generate better quality solutions. A thermal-aware mapping heuristic is proposed to generate some intelligent solutions, which become a part of the initial population in the

PSO. A trade-off has been established between communication cost and peak temperature of the die.

In [26], a reliability model which in turn considers thermal effect due to computation and communication entities present in the network, and also reliability MILP formulation of the mapping problem for improving system reliability with a constraint on the average packet delay of the network have been proposed. To address large sized applications, PSO based mapping method have been used. In [27], the author proposed an algorithm to intelligently perform a fault-tolerant resource allocation in real-time dynamic scenarios where tasks of applications are not known a priori. The slack times of the incoming tasks have been exploited in the application mapping/scheduling phase of the algorithm, to assign a fault tolerant policy to the corresponding task for mitigating the effect of transient faults. This helps to improve the deadline satisfaction of the task and also reduce the energy consumption. The proposed algorithm achieves 19.8%, 43.5% and 85.8% improvement in deadline satisfaction and on the average, the energy consumption is reduced by 29.1% and 6.7%, compared to other research. In [28], the author proposed a mapping methodology that reduces power consumption by decreasing the energy consumption in communication while guaranteeing the required performance providing an energy savings about 51%.

Most of the design-time methodologies adopt search-based approaches (e.g., GA, ILP, SA) that incur high computational costs. Thus, the evaluation time might not be acceptable for large scale problems. However, they provide efficient mapping solutions for small scale systems within acceptable time. The evaluation time can be reduced by efficient pruning of the search space, but at the risk of missing the highest quality mapping solutions. The reliability-aware design-time methodologies increase the system life-time, but they cannot overcome the faults incurred in the system.

B. Run-time mapping

In contrast to the design-time mapping, run-time mapping needs to account for the time taken to map each task as it contributes to overall application execution time. Therefore, typically greedy heuristic algorithms are used for efficient run-time mapping in order to optimize performance metrics. The need of run-time mapping is when the insertion of a new application into the system, which needs resources from the already executing applications, modifying parameters of a running application, killing a running application in order to free its occupied resources, changing performance requirements of a running application, or when current mapping is not sufficiently optimal, it requires (re-)mapping.

At run-time, the mapping of new applications to be supported onto a platform can be handled either by performing all the processing at the same time. The run-time platform manager handles the mapping of applications by taking the updated resources' status (Current System Status) into account. For on-the-fly mapping, efficient heuristics are required to assign new arriving tasks on the platform resources. These heuristics cannot guarantee for schedulability. However, these heuristics are platform independent since they do not use any platform specific analysis results computed in advance. Such heuristics lend well to map unknown applications (not available at design-time) on any platform.

In [29], the author investigates the performance of dynamic task mapping heuristics in NoC-base MPSoCs, targeting NoC congestion minimization. Tasks are mapped on demand, according to the NoC channels load. Results using congestion-aware mapping heuristics compared to a straight-forwardly defined heuristic achieve better results. In average, it is possible to reach up to 31% smaller channel load, up to 22% smaller packet latency, and up to 88% less.

In [30], dynamic master/slave applications on homogeneous NoC architectures have been investigated. Such applications have the characteristic that their program structure may be adapted at run time by inserting or removing tasks to react to changing requirements that are not known a priori. A heuristic is presented for an energy-and performance-aware assignment of dynamically created tasks to the processing units. The provided experimental results give evidence of the benefits of the proposed methods for synthetic test-cases as well as an adaptive image processing application.

In [31], the author demonstrates the designing of reconfigurable NoC-based MPSoC and programming it for real-life applications. The NoC is reconfigured at run-time to support different combinations of multiple applications at different times. Based on the investigations to map the applications on a 3×3 platform, the NoC reconfiguration overhead is kept at a minimum and the platform utilization is about 85%.

In [32], a number of communication-aware run-time mapping heuristics for the efficient mapping of multiple applications onto an MPSoC platform have been proposed in which more than one task can be supported by each processing element (PE). The proposed mapping heuristics examine the available resources prior to recommending the adjacent communicating tasks on to the same PE. In addition, the proposed heuristics give priority to the tasks of an application in close proximity so as to further minimize the communication overhead. The investigations show that the proposed heuristics are capable of alleviating NoC congestion bottlenecks when compared to existing alternatives. a map tasks of applications onto an 8x8 NoC-based MPSoC to show that the mapping heuristics consistently leads to reduction in the total execution time, energy consumption, average channel load and latency. In particular, an energy saved about 44% and average channel load is improved by 10% for some cases.

In [33], the author proposed a heuristic that illustrates how to incorporate the optimization factors for mapping multiple tasks onto MPSoC platforms, where communication takes place through a NoC. The heuristic attempts to balance the processing load on the platform PEs while reducing communication overhead by mapping highly communicating tasks on the same PE. For MPEG-4 application, the proposed technique reduces total execution time by 33%, resource usage by 37% and energy consumption by 40% when compared to state-of-the-art run-time mapping techniques.

In [34], the author presented Heterogeneous Near Convex Region Algorithm (HNCR), an incremental run-time mapping algorithm for heterogeneous NoC while adjusting the idea of

near convex region to fit the need of heterogeneous mapping. Experimental results show that HNCR gains 16.7% and 60.5% average reductions in network latency compared to greedy and random solutions, together with 17.9% and 54.9% average reductions in communication energy only at the cost of a slight increase in total execution time. In [35], the author proposed a distributed stochastic dynamic task mapping strategy for mapping applications efficiently onto a large dynamically reconfigurable NoC. The effectiveness of the mapping scheme was investigated considering the transient and steady states of the dynamic platform. The comparison with state-of-the-art centralized dynamic task mapping methods shows more than 26.4% improvement in application communication distance during steady state, which implies lower energy consumption and lower execution time.

In [36], the author proposed a proactive region selection strategy which prioritizes nodes that offer lower congestion and dispersion. The proposed strategy named: MapPro, quantitatively represents the propagated impact of spatial availability and dispersion on the network with every new mapped application in order to identify a suitable region to accommodate an incoming application that results in minimal congestion and dispersion. The simulation over different traffic patterns and observed gains of up to 41% in energy efficiency, 28% in congestion and 21% dispersion when compared to the state-of-the-art region selection methods.

In [37], The author addressed the optimization of NoC performances in term of power and latency, that is considered as a hybrid based-scheduling algorithm which evolve the cellular automata with genetic algorithm to solve mapping and scheduling problems, where each transition rule represents a chromosome allowing an automatically programming of the transition rules of the evolutionary cellular automata.

In [38], the author proposed a novel adaptive approach capable of handling dynamism of a set of applications on NoC. The applications are subject to throughput or energy consumption constraints. For each application, a set of non-dominated (Pareto) schedules are computed at design-time in the (energy, period, processors) space for different cores topologies. Then, upon the starting or ending of an application, a lightweight adaptive run-time scheduler reconfigures the mapping of the live applications according to the available resources. This run-time scheduler selects the best topology for each application and maps them to NoC. This scheduling approach is adaptive, it changes the mapping of applications during their execution, and thus delivers just enough power to achieve applications constraints.

C. Upcoming trends and open challenges in hybrid mapping

The reported mapping methodologies provide three alternatives: design-time mapping, on-the-fly mapping and hybrid (design-time analysis and then run-time use) mapping. Design-time techniques have pre-dominated the reported literature. However, their inability to handle dynamic workload scenarios has led to the formulation of on-the-fly mapping methodologies. On-the-fly strategies surmount the limitation of handling dynamic workloads at run-time, but with the fallout of possible non-optimal mapping due to limited compute power at run-time. Recently, the issues of design time and on-the-fly strategies have been addressed by developing hybrid mapping methodologies that attempt to incorporate the advantages of both. Hybrid strategies combine design space exploration of design-time techniques with the run-time management in order to select mapping configurations that are best suited to newly arriving applications.

They involve minimum computation at run-time, facilitating for light-weight run-time manager performing efficient mapping. The experimental results shows that runtime mapping gets speeded up by 93% when compared to state-of-the-art on-the-fly mapping methodologies [39]. Although the advantages of hybrid strategy seem promising, it comes with its own trade-offs due to the inherent pseudo dynamic nature and inability to handle new applications without available design-time exploration.

With no doubt, hybrid strategies seem to be followed in the field of mapping methodologies, but due to their nascent development and lack of in-depth examination, further development of design-time and on-the-fly mapping methodologies will continue hand-in-hand with hybrid strategies.

IV. CONCLUSION

In this paper, we presented a state of the art of most knowing mapping and scheduling technique on both static and dynamic. Developing dynamic or quasi-static mapping and scheduling algorithms for NoC could be another direction to explore. Run-time mapping theoretically exploit better the actual configuration and thus might improve the quality of mapping and scheduling. However, the cost of deciding run-time mapping must be justified by the quality improvement.

REFERENCES

[1] J. Flich, D. Bertozzi, "Designing Network On-Chip Architectures in the Nanoscale Era", Book Chapman & Hall/CRC, (2010).

[2] S. Kumar, A. Jantsch, J.-P.Soininen, M. Forsell, M. Millberg, J. Oberg, K. Tiensyrja, and A. Hermani, "A New on Chip Architecture and Design Methodology", in Proc. IEEE Computer Society Annual Symposium on VLSI, Apr. 25-26, pp. 105-112 (2002).

[3] J. Hu and R. Marculescu. Energy-aware mapping for tile-based NoC architectures under performance constraints. In Asia & South Pacific Design Automation Conference, Jan. (2003).

[4] S. Murali and G. D. Micheli. Bandwidth-constrained mapping of cores onto NoC architectures. In Design, Automation, and Test in Europe, pages 896–901. IEEE Computer Society, Feb. (2004).

[5] M. R. Garey and D. S. Johnson. Computers and intractability: a guide to the theory of NP-completeness.Freeman and Company, (1979).

[6] M. Sacanamboy, F. Bolaños, and R. Nieto, "A Primer for Mapping Techniques on NoC Systems", Embedded Systems and Applications the WorldComp International Conference Proceedings. Dec. (2014).

[7] Z. Huiyang, T. M. Conte, " Performance Modeling of Memory Latency Hiding Techniques", (2002).

[8] L. Liang-Yu,W. Cheng-Yeh, H. Pao-Jui, C. Chih-Chieh,J. Jing-Yang, "Communication-driven task binding for multiprocessor with latency insensitive network-on-chip", Asia and South Pacific Design Automation Conference, pp. 39-44, Jan. (2005).

[9] H. Hu, M. Radu,"Energy- and Performance-Aware Mapping for Regular NoC Architectures", IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, VOL. 24, NO. 4, pp. 551- 562, APRIL (2005).

[10] M. Ruggiero et al. Communication-aware allocation and scheduling framework for stream-oriented multi-processor systems-on-chip. In DATE, pages 3–8, (2006).

[11] N. Satish, K. Ravindran, and K. Keutzer. "A decomposition-based constraint optimization approach for statically scheduling task graphs with communication delays to multiprocessors". In DATE, pages 57–62, (2007).

[12] L. Thiele, I. Bacivarov, W. Haid, and K. Huang. Mapping Applications to Tiled Multiprocessor Embedded Systems. In ACSD, pp. 29–40, (2007).

[13] O. Heikki, K. Kangas, S. Erno, D.H. Timo, H. Marko,"", Journal of Systems Architecture: the EUROMICRO Journal, Vol. 53, No. 11, pp. 795-815, Nov. (2007).

[14] 1G. Chen, F. Li, S. Son, and M. Kandemir. "Application mapping for chip multiprocessors", In DAC, pp. 620–625, (2008).

[15] H. Javaid and S. Parameswaran. A design flow for application specific heterogeneous pipelined multiprocessor systems. In DAC, pages 250–253, (2009).

[16] B. H. Meyer, A. S. Hartman, and D. E. Thomas. Cost-effective slack allocation for lifetime improvement in noc-based mpsocs. In DATE, pp. 1596–1601, (2010).

[17] A. Bonfietti, L. Benini, M. Lombardi, and M. Milano. "An efficient and complete approach for throughput-maximal sdf allocation and scheduling on multi-core platforms", In DATE, pp. 897–902, (2010).

[18] L. Thiele, L. Schor, H. Yang, and I. Bacivarov. "Thermal-aware system analysis and software synthesis for embedded multi-processors", In DAC, pp. 268–273, (2011).

[19] W. Che and K. S. Chatha. "Unrolling and retiming of stream applications onto embedded multicore processors", In DAC, pp. 1272–1277, (2012).

[20] O. He, D. Sheqin, J. Wooyoung, B. Jinian, David Z. Pan," UNISM: Unified Scheduling and Mapping for General Networks on Chip", IEEE transactions on very large scale integration (VLSI) systems, Vol. 20, No. 8, pp. 1496-1509, Aug (2012).

[21] M. Norazizi Sham Mohd Sayuti, S. Indrusiak Leandro, "Real-time low-power task mapping in Networks-on-Chip", IEEE Computer Society Annual Symposium on VLSI (ISVLSI), (2013).

[22] S.K Pradip,M. Kanchan, S. Tapan, C. Santanu, "A Constructive Heuristic for Application Mapping onto Mesh Based Network-on-Chip", Journal of Circuits, Systems and Computers, Vol. 24, No. 08, (2015).

[23] F.Boutekkouk, "A Cellular Automaton Based Approach for Real Time Embedded Systems Scheduling Problem Resolution", on Artificial Intelligence Perspectives and Applications, Vol. 347, pp. 13-22, (2015).

[24] k.Sarzamin, A.Sheraz, A.G.Usman, M.Umer,F.Ishmanov, "An Efficient Algorithm for Mapping Real Time Embedded Applications on NoC Architecture", IEEE Access, Vol.6, pp. 16324 – 16335, (2018).

[25] M. Kanchan, M. Priyajit, C. Santanu, S. Sengupta, "Thermal-Aware Application Mapping Strategy for Network-on-Chip Based System Design", IEEE Transactions on Computers, pp. 528 – 542, Vol. 67, No. 4, Apr (2018).

[26] C. Navonil, M. Mukherjee, C. Chattopadhyay,"Reliability-aware application mapping onto mesh based Network-on-Chip", Int. J of Integration,Vol. 62, pp. 92-113, June (2018).

[27] C. Navonil, P. Suraj C. Santanu, "Task mapping and scheduling for network-on-chip based multi-core platform with transient faults" Journal of Systems Architecture, vol. 83, pp. 34-56, Feb. (2018).

[28] J. Hu and R. Marculescu, "Energy- and performance-aware mapping for regular NoC architectures", IEEE Trans. Comp.-Aided Des. Integ. Cir. Sys, Vol. 4, pp. 551–562, (2005).

[29] C. Ewerson, C. Ney, C. Fernando, "Investigating runtime task mapping for NoC-based multiprocessor SoCs", 17th IFIP International Conference on Very Large Scale Integration (VLSI-SoC), (2009).

[30] W. Stefan, Z. Tobias, T. Jürgen, "Run time mapping of adaptive applications onto homogeneous NoC-based reconfigurable architectures",International Conference on Field-Programmable Technology, (2009).

[31] K. S. Amit, K. Akash, S. Thambipillai, H. Yajun," Mapping Real-life Applications on Run-time Reconfigurable NoC-based MPSoC on FPGA", International Conference on Field-Programmable Technology, (2010).

[32] S. K Amit, S. Thambipillai, K. Akash, J. Wu Jigang, "Communication-aware heuristics for run-time task mapping on NoC-based MPSoC platforms", Journal of Systems Architecture, Vol. 56, pp.242–255,(2010).

[33] K. Samarth, S. Amit Kumar, S. Thambipillai, "Computation and communication aware run-time mapping for NoC-based MPSoC platforms", IEEE International SOC Conference, Sept. (2011).

[34] S. Jingcheng, T.Z. Chen, L. Li, "Incremental Run-time Application Mapping for Heterogeneous Network on Chip", IEEE 14th International Conference on High Performance Computing and Communication & IEEE 9th International Conference on Embedded Software and Systems, (2012).

[35] M. Hosseinabady,J.L. Nunez-Yanez, "Run-time stochastic task mapping on a large scale network-on-chip with dynamically reconfigurable tiles", ET Computers & Digital Techniques, Vol. 6, No. 1, pp. 1-11, Jan. (2012).

[36] H. Mohammad-Hashem, K. Anil, R. Amir-Mohammad, L.Pasi, J.Axel, T.Hannu "MapPro: Proactive Runtime Mapping for Dynamic Workloads by Quantifying Ripple Effect of Applications on Networks-on-Chip" IEEE/ACM International Symposium on Networks-on-Chip (NOCS),Jan. (2015).

[37] D.belkebir, F.Boutekkouk, "Two-steps into energy consumption optimisation due to the mapping of multimedia application to network on chip architecture",Int J of Intelligent Systems Technologies and applications,Vol. 15, No. 4, pp. 353-378,(2016).

[38] I. Assayad, A. Girault. "Adaptive Mapping for Multiple Applications on Parallel Architectures", Third International Symposium on Ubiquitous Networking, UNET'17, Casablanca, Morocco, May (2017).

[39] A. K. Singh, A. Kumar, and T. Srikanthan. A Hybrid Strategy for Mapping Multiple Throughput-constrained Applications on MPSoCs. In CASES, pages 175–184, 2011.

GPU-based Binary Particle Swarm Optimization For Bitmap Join Indexes Selection Problem In Data Warehouses

Lyazid TOUMI
Computer Science Department
College of Sciences, Ferhat Abbas
University Setif 1
Setif 19000, Algeria
Lyazid.toumi@univ-setif.dz

Ahmet UGUR
Computer Science Department
Central Michigan University
Mount Pleasant, 48859, MI, USA
Ahmet.ugur@cmich.edu

Yamina AZZI
Computer Science Department
College of Sciences, Ferhat Abbas
University Setif 1
Setif 19000, Algeria
azzimina14@gmail.com

***Abstract*—**

Data warehouses are very large databases and the crucial part of business intelligence. The performance of a data warehouse is an important aspect and its optimization is a difficult task. The emergence of the graphics processing unit (GPU) based computation in recent years has brought a potential in a range of scientific applications. The usage of GPUs in databases technologies, more precisely in the physical design phase, is a potential domain for optimization tasks. The Bitmap Join Indexes selection problem (BJISP) is crucial in the physical data warehouse design. In the present work, a GPU-based parallel binary particle swarm optimization (GBPSO) approach is proposed for solving the BJISP. Experiments have been performed to demonstrate the effectiveness of the proposed approach against the best serial approach for solving the BJISP. Furthermore, scalability experiments were performed to observe the behavior of the proposed approach against the best comparable approach which is serial in nature. In all experiments, the GBPSO is found to be considerably more effective than the best competitor algorithm.

Keywords*— *GPU, Parallel binary particle swarm optimization, Data warehouse, Query optimization, Bitmap join index, Bitmap join indexes selection problem.

I. INTRODUCTION AND BACKGROUNDS

The physical design in a data warehouse is an optimization problem known to be hard [1]. The physical design techniques in data warehouses can be classified into two main categories:

a) The redundant family, which utilize an extra memory space for the optimization techniques, (e.g. indexes and materialized views).

b) The non-redundant family, which do not utilize an extra space for the optimization techniques. (e.g. horizontal and vertical partitioning, referential horizontal partitioning and potential parallelism of these).

The data warehouse-based applications generally require a very complex online analytical processing (OLAP). The OLAP queries are computationally expensive and can take a very long time, because of the number of joins performed and expansion of volumes of data used along data cube dimensions. Join techniques, like, hash join, merge join, and nested-loop join become ineffective due to increase in both the number of joins and the information volume [5]. The single table indexation methods utilized in relational databases, for example, B-Tree, hash and bitmap indexes are limited in data warehouses [5,6,7]. Bitmap join index (BJI) introduced by O'Neil et al. allows the indexation of multiple tables [8]. The BJI allows to pre-calculate the joins among several tables [5, 8].

The index selection problem (ISP) in the redundant family physical database design is to choose a configuration of indexes to be created to optimize the cost of the queries. The ISP is an important problem largely tackled in the classical database physical design, in both centralized and distributed context [6,9,10,11,12,13]. In data warehouse context, the bitmap join index selection problem (BJISP) is an important problem for the data warehouse administrator (DWA), and the BJISP is much harder than ISP and known to be NP-hard [14]. The BJISP deals with $2^n - 1$ possibilities in the case of a single non-key attribute indexation and $2^{2^n} - 1$ possibilities in the case of the multiple non-key attributes indexation, n being the number of non-key attributes.

As stated earlier, the BJISP is a well-known optimization problem in data warehouse physical design, The BJISP is formalized as follows [1,2,3]:

- data warehouse DW with a set of dimension tables D and a fact table F,
- query workload Q, a set of queries defined on DW schema,
- non-key dimension attributes A extracted from Q,
- storage space constraint S.

The objective is to find a configuration C defined on non-key attributes in A such that the global cost of the query workload, called GlobalCost(Q, C), is minimized and the storage constraint S is satisfied. We have used the same global cost model as defined in Toumi et al. [3,4].

Several approaches to solve BJISP exists: approaches based on data mining techniques (DM) [14,15], approaches based on meta-heuristic methods such as genetic algorithm s (GA) [4,16], binary particle swarm optimization (BPSO) [4], and artificial immune system (AIS) [17], the minimal transversal based approach [18] and the linear programming based approach [3].

Toumi et al. demonstrated the effectiveness of the BPSO approach against the well-known approaches cited above [4]. But the main limitation of the serial particle swarm optimization (PSO) is the prolonged computation time to find the optimal/suboptimal solution for large-scale BJIS problem.

Several parallel PSO methods have been proposed [19–22] for solving some general purpose, non-database related optimization problems. In the present work, a GPU-based parallel BPSO is proposed for solving the BJISP. Several experiments were performed to demonstrate the effectiveness and advantages of the proposed approach against the best well-known method to solve BJISP so far.

The rest of this paper is organized as follows: Section 2 presents the details about the proposed method, Section 3 presents the experimental results, and conclusions are presented in Section 4.

II. PROPOSED METHOD

The optimization problems, like the BJISP can be easily solved by a distributed paradigm known as Swarm Intelligence (SI). Gerardo Beni et al. have introduced the SI inspired from the biological examples such as bird flocking, ant colonies, animal herding, fish schooling and bacterial growth [23]. The PSO is a population-based method based on swarm intelligence introduced by Eberhart et al. [24].

Algorithm Particle swarm optimization pseudo-code

1: $Swarm \leftarrow$ GenerateInitialSwarm(x_i^0);
2: $t \leftarrow 1$;
3: **while** t ≤ MAX **do**
4: $w_t \leftarrow Inertia(t)$ using Eq.5;
5: **for each** particle $x_i^{t-1} \in Swarm$ **do**
6: Evaluate particle x_i^{t-1};
7: **if** fitness (x_i^{t-1}) is better than fitness (p_i^{t-1}) **then**
8: $p_i \leftarrow x_i^{t-1}$;
9: **end if**
10: **if** fitness (p_i^{t-1}) is better than fitness (g) **then**
11: $p_g \leftarrow p_i^{t-1}$;
12: **end if**
13: $v_i^t \leftarrow Velocity(x_i^{t-1}, v_i^{t-1})$ using Eq. 1;
14: **if** $|v_i^t| > V_{max}$ **then**
15: clamp it to $|v_i^t| \leftarrow V_{max}$
16: **end if**
17: $x_i^t \leftarrow Location(v_i^t)$ using Eq. 2;
18: **end for**
19: $t \leftarrow t+1$;
20: **end while**

Fig. 1. Particle swarm optimization pseudo-code

Fig. 1 describes the PSO algorithm. The particle in the swarm is defined by its location x_i and its velocity v_i. The best location covered by the particle i is called p_i, and the best value of all p_i values is called p_g. On each generation, a particle's location and velocity are revised allowing its own p_i and p_g values. This behavior is described by Eqs. (1) and (2):

$$v_{ij}^t = w v_{ij}^{t-1} + c_1 r_1 (p_{ij} - x_{ij}^{t-1}) + c_2 r_2 (p_{gj} - x_{ij}^{t-1}) \quad (1)$$

$$x_{ij}^t = x_{ij}^{t-1} + v_{ij}^t \quad (2)$$

v_{ij}^{t-1} is the current velocity, v_{ij}^t is the new speed of particle i, w is the inertia weight, c_1, and c_2, are two positive constants, r_1, and r_2 are the uniformly distributed random numbers in [0,1], $x_{ij}^{t-1}$is the current location of particle i at time t and x_{ij}^t is new location of particle i. The velocity v_{ij} is bounded by the range of velocities $[-V_{max}, V_{max}]$ to prevent the particle from flying out of the solution space.

For solving problems in binary space, Kennedy and Eberhart have introduced an extended version of PSO called Binary Particle Swarm Optimization (BPSO) [25]. In the BPSO, a particle changes in a space limited to zero and one on each dimension (Eq. 3). A sigmoid transformation is applied to the velocity in Eq. (4) to compute the chance of the j[th] element value of the particle location taking the value 1. Velocities are updated as in Eq. (1), but locations are updated as in Eq. (3).

$$x_{ij}^t = \begin{cases} 1, if\ random < s(v_{ij}^t) \\ 0, \quad otherwise \end{cases} \quad (3)$$

$$s(v_{ij}^t) = \frac{1}{1+e^{-v_{ij}^t}} \quad (4)$$

We have used time-varying inertia weight described in Eq. (5) to improve the performance of the BPSO [26]:

$$w(t) = w_{max} - \frac{(w_{max} - w_{min})}{K} \times t \quad (5)$$

The fitness function is computed as follows [14, 3]:

$$\text{fitness(C)} = \begin{cases} \text{GlobalCost (Q, C)} \times 2^{pen(c)}, if\ pen(c) > 1 \\ \text{GlobalCost (Q, C)}, \ otherwise \end{cases} \quad (6)$$

$$\text{Pen(C)} = \frac{\text{GlobalStorage(c)}}{S} \quad (7)$$

The GPUs are used in a various fields of engineering and in simulation [23,27]. The architecture of the GPU framework is mainly composed of two hosts, the first one is from Central Processing Unit (CPU) and main memory, the second one is from several blocks of threads that access a shared memory.

The main motivation to use a GPU-based approach is that the GPU usually has hundreds of cores in comparison to about dozen of cores of CPU. The GPU allows tasks that can be efficiently divided across many threads. In the present work, a GPU-based parallel BPSO method (GBPSO) is proposed and described in Fig. 2. The CUDA parallel processing platform for the GPU (NVIDIA Corporation) is used to implement the proposed approach [28].

The fitness functions of the serial BPSO to solve BJISP require more computation time [4]. If each particle's fitness in the swarm can be computed independently (i.e., in parallel), the computation time will be decreased, which in turn, would

improve the efficiency of the BPSO method. The parallel computation of fitness can be exploited through GPU. Furthermore, both updating and validation of the velocity and location, computing of the personal and global best fitness can also done by GPU in parallel. The CPU (host) controls the initialization part *GenerateInitialSwarm* and the master operations in the BPSO code presented in Fig. 1. The rest of code is performed by the GPU device through the CUDA platform.

The CUDA platform has three main categories of functions [28]:

a) The Host Functions: called and performed only by the CPU.

b) The Kernel Functions: called only by CPU and performed by the GPU.

c) The Device Functions: called and performed only by the device (GPU).

Algorithm GPU-Based Parallel BPSO pseudo-code

1: GenerateInitialSwarm(x_i^0);
2: **for** t=1 to MAX **do**
3: Compute the fitness using the kernel function of FitnessKernel;
4: Update p_i and p_g;
5: Update x_i^{t+1} and v_i^{t+1};
6: **end for**
7: Return p_g

Fig. 2. GPU-Based Parallel BPSO (GBPSO) algorithm pseudo-code.

Algorithm FitnessKernel function pseudo-code

1: Transfer location vectors of all particles from CPU to GPU;
2: syncthreads();
3: Evaluate particle x_i^t; in the parallel reduction

Fig. 3. FitnessKernel function pseudo-code

GBPSO algorithm presented in Fig. 2 is implemented in the CUDA platform. Each particle is represented by one thread block, and each query in the query workload is presented by one thread. The number of thread blocks is equal to the number of particles in the swarm. Each thread calculates the fitness of one query. The number of threads in each thread block is equal to the number of queries in the query's workload. After the computing of the fitness of each query in the same block, the sum of the fitness of all queries of the particle is performed.

A pseudocode of the kernel function is shown in Fig. 3. The Fitnesskernel function is executed for each query in the query workload in parallel. This function transfers the location vectors of all particles to the global memory once at the start of each iteration. Each thread reads the location vector of one particle, from the global memory to the shared memory of its block, and only computes the query workload cost which has the same index as the thread.

III. EXPERIMENTAL RESULTS

A. Problem instances

In the experiments, the benchmark APB-I [29] is used to generate the data warehouse. The star schema of this data warehouse consists of four dimensions tables: CHANLEVEL (9 rows), CUSTLEVEL (900 Rows), PRODLEVEL (9,000 rows), TIMELEVEL (24 rows) and the fact table ACTVARS (24,786,000 rows). Three classes of problems used in the experiments are the flowing: (see [4])

- Smaller size problem set (CSP): this class contains 100 OLAP queries and use at most 12 non-key attributes from dimension tables: {all, year, retailer, quarter, month, line, group, family, division, class, gender, city} with cardinalities: 9, 2, 99, 4, 12, 15, 300, 75, 4, 605, 2 and 255, respectively.
- Moderate size problem set (CMP): this class consists of 250 OLAP queries and use at most 16 non-key attributes from dimension tables: {division, line, family, group, class, status, year, quarter, month, day, state, city, retailer, type, gender, all} with cardinalities: 4, 15, 75, 300, 605, 5, 2, 4, 12, 5, 45, 255, 99, 10, 2 and 9, respectively.
- Larger size problem set (CLP): this class consists of 500 OLAP queries and use at most 20 non-key attributes from dimension tables: {all, year, retailer, quarter, month, day, line, group, family, division, class, gender, city, state, type, educational, marital, supplier, status, category} with cardinalities: 9, 2, 99, 4, 12, 5, 15, 300, 75, 4, 605, 2, 255, 45, 10, 6, 4, 15, 5 and 3, respectively.

B. Performance evaluation

A set of experiments were performed to analyze the efficiency of the GBPSO against the best serial method for solving BJISP using the three problem classes CSP, CMP and CLP described above. The best serial method for solving BJISP is shown to be the serial version of the same algorithm, BPSO [4] All tests are performed under Bi-CPU Intel Xeon E5-2680v4 (28-core) with 256 GB RAM, and Nvidia Tesla K80 GPUs (4992-cores) with 24 GB dedicated Memory. For both GBPSO and BPSO, the parameters c_1, c_2, V_{max}, w_{max}, w_{min} were set to 2.0, 2.0, 6.0, 0.95 and 0.5 respectively. The population size and the number of iterations were set to 30 and 2000 respectively yielding 60000 maximum evaluations (2000 x 30).

Tables I, II and III. demonstrate the number of disk page accesses needed (I/O costs) in order to execute the query workload for the problem sets CSP, CMP and CLP for sixteen different storage capacity sizes. The column Value represents the value of the best minimum cost solution found. The column Time shows the computation time in seconds. In each table row the best (i.e., the minimum) value and fastest computation time were presented in bold for both of the algorithms considered.

1) The smaller size problem set CSP results

The querying performance for the smallest size problem CSP presented in Table I demonstrates that the GPSO approach generates better results in general. The GBPSO has outperformed the BPSO approach for both best optimal/sub-optimal solutions found and the computation time. The GPSO

approach has generated best solutions in all 16 cases (100 %), while the BPSO algorithm has generated best solutions in 13 out of 16 cases (81.25%). In terms of the computation time, the GBPSO was about 2 (1.72 exactly) times faster than the BPSO. In summary, as also indicated by the last row of Table I., the GBPSO approach has shown much improved performance than BPSO approach.

TABLE I. QUERYING PERFORMANCE RESULTS FOR THE SMALLER SIZE PROBLEM SET CSP.

S (MB)	BPSO		GBPSO	
	Value	Time	Value	Time
500	**44,558,320.52**	2.30	**44,558,320.52**	**1.20**
600	**44,558,320.52**	2.20	**44,558,320.52**	**1.10**
700	**44,558,320.52**	2.20	**44,558,320.52**	**1.20**
800	44,558,320.52	2.20	**41,647,924.96**	**1.20**
900	**41,647,924.96**	2.40	**41,647,924.96**	**2.40**
1000	**41,647,924.96**	2.30	**41,647,924.96**	**1.10**
1100	**41,595,428.61**	2.40	**41,595,428.61**	**1.20**
1200	41,485,344.25	2.40	**40,752,322.46**	**1.10**
1300	**39,017,671.99**	2.40	**39,017,671.99**	**1.10**
1400	**39,017,671.99**	2.50	**39,017,671.99**	**1.10**
1500	38,685,033.05	**2.50**	**38,574,948.69**	2.60
1600	**38,574,948.69**	2.50	**38,574,948.69**	**2.50**
1700	**36,443,236.53**	2.70	**36,443,236.53**	**1.20**
1800	**36,443,236.53**	2.60	**36,443,236.53**	**1.20**
1900	**36,054,780.08**	2.60	**36,054,780.08**	**1.30**
2000	**35,944,695.72**	2.70	**35,944,695.72**	**1.10**
AVG	40299448.7	2.43	**40064854.9**	**1.41**

TABLE II. QUERYING PERFORMANCE RESULTS FOR THE MODERATE SIZE PROBLEM SET CMP.

S (MB)	BPSO		GBPSO	
	Value	Time	Value	Time
500	**148,060,488.95**	10.10	**148,060,488.95**	**2.80**
600	**148,060,488.95**	12.40	**148,060,488.95**	**2.70**
700	**148,060,488.95**	11.00	**148,060,488.95**	**2.80**
800	**148,060,488.95**	12.40	**141,120,480.53**	**2.80**
900	**141,120,869.53**	12.30	**141,120,869.53**	**2.70**
1000	**141,064,739.78**	11.90	141,120,869.53	1.30
1100	**141,120,480.53**	12.60	**141,120,480.53**	**2.80**
1200	140,294,240.08	11.70	**135,131,781.39**	**2.80**
1300	**134,180,861.11**	11.20	**134,180,861.11**	**2.80**
1400	**134,124,731.36**	9.80	134,180,861.11	**2.80**
1500	134,349,390.03	9.50	**134,180,861.11**	**2.80**
1600	133,298,490.91	9.50	**131,437,615.13**	**2.80**
1700	128,361,079.90	10.0	**127,354,029.86**	**2.70**
1800	127,968,325.95	10.50	**127,525,681.39**	**2.70**
1900	**127,525,292.39**	9.10	**127,525,292.39**	**2.80**
2000	127,967,936.95	9.60	**127,581,422.14**	**2.80**
AVG	137,726,149.65	10.85	**136,735,160.79**	**2.68**

2) The moderate size problem set CMP results

The querying performance for the moderate size problem CMP presented in Table II indicates that the GPSO approach generates better results in general. For both best solutions found and computation time, the GBPSO approach has outperformed the BPSO approach. The GBPSO approach has produced best solutions in 14 out of 16 cases (87.5%) while the BPSO approach has generated best solutions in 10 out of 16 cases (62.5%). As far as the computation time, the GBPSO was around 4 (4.05 precisely) times faster than the BPSO. In summary, as also indicated by the last row of Table II., the GBPSO algorithm has shown considerably better performance than the BPSO method in all aspects.

3) The larger size problem set CLP results

Table III shows the querying performance for the GBPSO and BPSO approaches for the larger size problem set CLP. Note that this class is the hardest one. The BPSO algorithm has again generated better results in general. The GBPSO method has outperformed the BPSO method for both best solutions found and average solution quality obtained. The GPSO method has generated best solutions in all 16 cases (100 %) while the BPSO method has generated best solutions in 7 out of 16 cases (43.75 %). In terms of the computation time, the GBPSO method was about 6 times (6.31 exactly) faster than the BPSO method. In summary, as also indicated by the last row of Table III., the GBPSO method has again shown considerably better performance than the BPSO method in all aspects.

TABLE III. QUERYING PERFORMANCE RESULTS FOR THE LARGER SIZE PROBLEM SET CLP.

S (MB)	BPSO		GBPSO	
	Value	Time	Value	Time
500	**356,324,650.47**	17.70	**356,324,650.47**	**3.60**
600	**356,324,650.47**	17.50	**356,324,650.47**	**1.70**
700	**356,324,650.47**	17.80	**356,324,650.47**	**1.70**
800	356,324,650.47	17.70	**353,415,287.32**	**3.60**
900	**341,269,799.59**	18.40	**341,269,799.59**	**1.70**
1000	**341,269,799.59**	18.50	**341,269,799.59**	**3.60**
1100	**341,269,799.59**	18.20	**341,269,799.59**	**2.80**
1200	341,269,799.59	18.40	**331,642,486.63**	**3.60**
1300	**327,446,529.75**	18.90	**327,446,529.75**	**4.80**
1400	327,343,869.80	19.00	**326,702,600.11**	**3.70**
1500	327,237,990.10	19.00	**325,030,648.71**	**3.50**
1600	327,237,990.10	18.90	**319,271,103.97**	**3.50**
1700	314,963,508.52	19.60	**312,866,684.24**	**3.40**
1800	313,735,637.98	19.90	**313,520,599.97**	**1.80**
1900	313,520,599.96	19.70	**312,671,020.20**	**2.60**
2000	313,414,720.26	19.80	**313,312,060.32**	**1.70**
AVG	334704915.42	18.69	**333041398.21**	**2.96**

C. Scalabality study

Experiments were expanded (i.e., scaled up) to further analyze the efficiency of the GBPSO method against the BPSO method. The cost model used in the previous experiments was also used here. In the scalability study, the fact table size has been expanded from 30 million to 150 million tuples (in 30 million tuple increments, yielding a total of 5 different cases) and for every fact table size |F|, the storage capacity S was expanded from 500 to 2,000 MB (in 500 MB increments, yielding a total of 4 different cases).

The querying performance of scalability experiments for the two problem sets CMP and CLP are presented in Tables IV. and V. (see Sect. III.B for table details). The additional column |F| in the tables represents the fact table size in millions.

1) Scallability results for the moderate size problem CMP

The performance of scalability experiments for the moderate size problem set CMP presented in Table IV indicates that the GBPSO algorithm has again generated better results than the

BPSO algorithm in general. The GBPSO has outperformed the BPSO algorithm for both best solutions found and computation time. The GBPSO has generated best solutions in 19 out of 20 cases (95%), almost in all runs except in the case when the storage size S was equal to 1,500 and the fact table size, |F|, was equal to 60M. The BPSO approach has generated best solutions in 16 out of 20 runs (80%). In terms of the computation time, the GBPSO was about 6 times (6.15 exactly) faster than the BPSO. In summary, as also indicated by the last row of Table IV., the GBPSO algorithm has again shown considerably better performance than the BPSO approach in all aspects.

TABLE IV. SCALABILITY RESULTS FOR THE MODERATE SIZE PROBLEM SET CMP.

S (MB)	|F|	BPSO		GBPSO	
		Value	Time	Value	Time
500	30	**179,204,786.30**	7.30	**179,204,786.30**	**1.40**
	60	**384,146,490.57**	12.80	**384,146,490.57**	**1.30**
	90	**576,214,972.02**	7.10	**576,214,972.02**	**1.30**
	120	**768,283,453.46**	8.00	**768,283,453.46**	**1.30**
	150	**960,351,934.91**	7.90	**960,351,934.91**	**1.40**
1000	30	188,846,319.50	7.30	**170,804,708.72**	**1.30**
	60	**358,400,973.58**	8.00	**358,400,973.58**	**1.30**
	90	**576,214,972.01**	6.80	**576,214,972.01**	**1.20**
	120	**768,283,453.46**	7.60	**768,283,453.46**	**1.30**
	150	**960,351,934.91**	7.20	**960,351,934.91**	**1.30**
1500	30	163,080,934.29	7.70	**162,608,876.40**	**1.30**
	60	**358,400,973.58**	9.00	367,210,306.89	**1.30**
	90	**537,597,160.86**	7.10	**537,597,160.86**	**1.30**
	120	**768,283,453.46**	8.00	**768,283,453.46**	**1.30**
	150	**960,351,934.91**	8.30	**960,351,934.91**	**1.30**
2000	30	159,084,809.74	8.20	**159,018,521.38**	**1.30**
	60	342,953,831.21	9.60	**342,007,547.35**	**1.30**
	90	**537,597,160.86**	7.20	**537,597,160.86**	**1.30**
	120	**716,793,348.14**	8.30	**716,793,348.14**	**1.30**
	150	**960,351,934.91**	7.80	**960,351,934.91**	**1.30**
AVG		561,239,741.63	8.06	**560,703,896.26**	**1.31**

2) Scallability results for the larger size problem CLP

Table V illustrates the performance of the GBPSO algorithm against the BPSO algorithm for the larger size problem CLP (the hardest class) in scalability. The GBPSO algorithm again generated better results than those generated by the BPSO algorithm. The GBPSO has outperformed the BPSO algorithm for both best solutions found and computation time. The GBPSO has always generated best solutions in all runs (100%), while the BPSO approach has generated best solutions in 16 out of 20 runs (80%). In terms of the computation time, the GBPSO was about 6 times (10.28 exactly) faster than the BPSO. In summary, as also indicated by the last row of Table V., the GBPSO algorithm has again shown considerably better performance than the BPSO approach in all aspects.

TABLE V. SCALABILITY RESULTS FOR THE LARGER SIZE PROBLEM SET CLP.

S (MB)	|F|	BPSO		GBPSO	
		Value	Time	Value	Time
500	30	**431,277,020.79**	17.60	**431,277,020.79**	**2.30**
	60	**909,959,419.24**	16.50	**909,959,419.24**	**1.80**
	90	**1,364,927,986.13**	24.50	**1,364,927,986.13**	**1.70**
	120	**1,819,896,553.03**	16.80	**1,819,896,553.03**	**1.70**
	150	**2,274,865,119.92**	17.70	**2,274,865,119.92**	**1.70**
1000	30	416,168,229.91	18.00	**415,087,005.29**	**1.80**
	60	**862,533,466.83**	17.50	**862,533,466.83**	**1.80**
	90	**1,364,927,986.12**	19.40	**1,364,927,986.12**	**1.70**
	120	**1,819,896,553.03**	17.10	**1,819,896,553.03**	**1.80**
	150	**2,274,865,119.92**	16.60	**2,274,865,119.92**	**1.70**
1500	30	401,061,640.29	18.60	**396,323,138.85**	**1.70**
	60	**862,533,466.82**	17.50	**862,533,466.82**	**1.70**
	90	**1,293,789,912.87**	20.50	**1,293,789,912.87**	**2.10**
	120	**1,819,896,553.03**	17.10	**1,819,896,553.03**	**1.70**
	150	**2,274,865,119.92**	16.80	**2,274,865,119.92**	**1.70**
2000	30	386,090,237.72	18.80	**385,008,644.29**	**1.70**
	60	832,316,680.21	18.00	**829,067,899.00**	**1.70**
	90	**1,293,789,912.87**	21.00	**1,293,789,912.87**	**1.70**
	120	**1,725,046,358.90**	17.50	**172,5046,358.90**	**1.70**
	150	**2,274,865,119.92**	16.70	**2,274,865,119.92**	**1.70**
AVG		1,335,178,622.87	18.21	**1,334,671,117.84**	**1.77**

IV. CONCLUSIONS

In the present work, we have proposed a GPU-based parallel binary particle swarm optimization approach GBPSO to solve the bitmap join indexes selection problem BJISP in data warehouses. To prove the effectiveness of the proposed GBPSO approach, we have utilized several classes of problem sets, the smaller size, the moderate size and the larger size. Several experiments have been performed to validate our approach against the best serial comparable approach in the literature, the binary particle swarm optimization, BPSO, on a benchmark data warehouse (APB-1). We have also executed the GBPSO and BPSO approaches to further investigate the performance in scalability by systematically increasing both the fact table size and storage size. Both the general and scalability results have demonstrated that the GBPSO approach has outperformed the BPSO method in numerous aspects for solving the BJISP. The GBPSO method has found the optimal/suboptimal solutions significantly faster than the BPSO method for all larger problem sets considered (CMP and CLP).

The GBPSO could be utilized in other data warehouses physical design problems such as vertical partitioning, referential horizontal partitioning, and materialized view selection.

REFERENCES

[1] K. Aouiche, O. Boussaid, F. Bentayeb. "Automatic selection of bitmap join indexes in data warehouses." In: Proceedings of international conference on data warehousing and knowledge discovery, vol. ,pp 64–73, 2005.

[2] L. Bellatreche, R. Missaoui, H. Necir et al., "A data mining approach for selecting bitmap join indices." J Comput Sci Eng, vol. 1(2), pp.206–223, 2008.

[3] L Toumi, A. Moussaoui, A. Ugur, "A linear programming approach for bitmap join indexes selection in data warehouses", Procedia Computer Science, vol. 52(c), pp. 169-177 , 2015.

[4] L. Toumi, A. Moussaoui, A. Ugur, "Particle swarm optimization for bitmap join indexes selection problem in data warehouses." The Journal of Supercomputing, vol. 68(2), pp. 672–708, 2015.

[5] P. O'Neil, D. Quass, "Improved query performance with variant indexes." In: Proceedings of the ACM SIGMOD international conference on management of data, pp 38–49, 1997.

[6] S. Agrawal, S. Chaudhuri, VR. Narasayya "Automated selection of materialized views and indexes in sql databases." VLDB conference, pp 496–505, 2000.

[7] DC .Zilio, J. Rao, S. Lightstone, G. Lohman, A. Storm, C. Garcia-Arellano "Db2 design advisor: integrated automatic physical database design." In: Proceedings of the Thirtieth international conference on Very large databases, vol. 30, pp.1087–1097, 2004.

[8] P. O'Neil, G. Graefe "Multi-table joins through bitmapped join indices." ACM SIGMOD Record vol. 24(3), pp 8–11, 1995.

[9] J. Kratica, I. Ljubic, D. Tosic "A genetic algorithm for the index selection problem". Springer, 2003.

[10] D. Comer "The difficulty of optimum index selection.", ACM Transactions on Database Systems (TODS), vol. 3(4), pp. 440–445, 1978.

[11] M.T. Ozsu, "Principles of Distributed Database Systems.", Prentice Hall Press; 3rd ed., 2007.

[12] S. Chaudhuri,M. Datar, V. Narasayya "Index selection for databases: A hardness study and a principled heuristic solution.", Knowledge and Data Engineering, IEEE Transactions, vol. 16(11), pp. 1313–1323, 2004.

[13] A. Caprara, M. Fischetti, D. Maio "Exact and approximate algorithms for the index selection problem in physical database design.", Knowledge and Data Engineering, IEEE Transactions, vol. 7(6), pp. 955–967, 1995.

[14] K. Aouiche, J. Darmont, O. Boussaid, F. Bentayeb "Automatic selection of bitmap join indexes in data warehouses." In: DataWarehousing and Knowledge Discovery, Springer, pp. 64–73, 2005.

[15] L. Bellatreche, R. Missaoui, H. Necir, H. Drias "A data mining approach for selecting bitmap join indices." JCSE, vol. 1(2), pp. 177–194, 2007.

[16] R. Bouchakri, L. Bellatreche "On simplifying integrated physical database design." In: Advances in Databases and Information Systems. Springer, pp. 333–346, 2011.

[17] A. Gacem, K. Boukhalfa "Immune algorithm for bitmap join indexes." In: Neural Information Processing. Springer, pp. 560–567, 2012.

[18] I. Ghabry, S.B. Yahia, M.N. Jelassi "Selection of Bitmap Join Index: Approach Based on Minimal Transversals". In: Ordonez C., Bellatreche L. (eds) Big Data Analytics and Knowledge Discovery. DaWaK 2018. Lecture Notes in Computer Science, vol 11031, Springer, 2018.

[19] L. Mussi, F. Daolio, and S. Cagnoni, "Evaluation of parallel particle swarm optimization algorithms within the CUDA_architecture," Information Sciences, vol. 181, no. 20, pp. 4642–4657, 2011.

[20] L. De P. Veronese and R. A. Krohling, "Swarm's flight: Accelerating the particles using C-CUDA," in Proceedings of the 2009 IEEE Congress on Evolutionary Computation, CEC 2009, pp. 3264–3270, nor, May 2009.

[21] W. Wang, Y. Hong, and T. Kou, "Performance gains in parallel particle swarm optimization via NVIDIA GPU," in Proceedings of theWorkshop on ComputationalMathematics andMechanics, 2009.

[22] Y. Zhou and Y. Tan, "GPU-based parallel particle swarm optimization," in Proceedings of the IEEE Congress on Evolutionary Computation (CEC '09), pp. 1493–1500, May 2009.

[23] J. Wang and G. Beni, "Object Analysis of Multi-Valued Images," 11-July1989, uS Patent 4,847,786.

[24] J. Kennedy, R. Eberhart "Particle swarm optimization." In: Proceedings of the IEEE international conference on neural networks, pp 1942–1948, 1995.

[25] J. Kennedy, R.C. Eberhart "A discrete binary version of the particle swarm algorithm." In: Proceedings of IEEE international conference on systems, man, and cybernetics, pp 4104–4108,1997.

[26] J. Xin, G. Chen, and Y. Hai., A Particle Swarm Optimizer with Multistage Linearly-Decreasing Inertia Weight, In Computational Sciences and Optimization, 2009. CSO 2009. International Joint Conference on, volume 1, pages 505–508. IEEE, 2009.

[27] D. Luebke et al., "General-purpose Computation on Graphics Hardware," in Workshop, SIGGRAPH, 2004.

[28] NVIDIA, CUDA programming guide v.10.1, NVIDIA Corporation, https://docs.nvidia.com/cuda/cuda-c-programmingguide/, 2019.

[29] APB-I, OLAP Benchmark, Release II, OLAP Council. http://www.olapcouncil.org/, 1998.

Multi-objective modeling for the integrated production and distribution planning: Cost vs. Energy

Besma ZEDDAM
Manufacturing Engineering Laboratory of Tlemcen, University of Tlemcen
Tlemcen, Algeria
b.zeddam@hotmail.com

Fayçal BELKAID
Manufacturing Engineering Laboratory of Tlemcen, University of Tlemcen
Tlemcen, Algeria
f_belkaid@yahoo.fr

Mohammed BENNEKROUF
ESSA Tlemcen, Algeria
Tlemcen, Algeria
mbenkrouf@yahoo.fr

***Abstract*—integrated planning is becoming the most dominant over the operational research field because of its efficiency and its ability to cover the different aspects of the problem. Production routing problem is one of these problems of the integrated planning that interests to jointly optimize production, inventory and distribution planning. This paper has the purpose of developing two mono-objective models for the Production-Routing Problem, one of them minimizes the total costs which is the classical problem but with an energy capacity constraint, while the other one minimizes the energy consumed by the production system. Finally, a multi-objective model is proposed to combine the two objectives mentioned previously using *LP-metric* method in the context of sustainable supply chain, computational results are also presented and discussed through the different scenarios.**

Keywords—Production-Routing Problem (PRP), energy consumption, energy capacity constraints, multi-objective, LP-Metric.

I. INTRODUCTION

Supply chain optimization is a huge field that aims to the best organization of the companies' services of the whole process from the first supplier to the final customer where there are so many issues to deal with. Recently the focus on the "integrated supply chain" has become bigger in fact that its benefits has been proved considering the different aspects treated using different methods.

The Production Routing Problem (PRP) is one among these integrated problems, its main objective is to minimize the total cost that include production, setup, inventory and transportation costs. It is a difficult problem that aim to jointly optimize production, inventory and distribution decisions.

The PRP is a complicated problem that combines two well-known classical problems that's have been extensively studied that are the Lot-Sizing Problem (LSP) and the Vehicle Routing Problem (VRP) [1].The lot sizing problem is the problem of determining through a given planning horizon the best production planning with the appropriate decisions of the amount of product (or products) to produce and to store according to a set of demands, while the VRP serves to find the optimal routes for the vehicle.

This paper introduces a new approach to solve the Production-Routing Problem, we first consider the classical objective which is the minimization of the total costs, then we treat the problem with energy consideration, i.e. minimizing the energy consumption of the production system, and finally we joint both objectives in only one using LP-metric method which is one of the multi-objectivity methods.

II. LITERATURE REVIEW

The production routing problem has been widely studied since it was introduced by Chandra [2] where the considered objective was to minimize the total costs that are production, setup, inventory and transportation costs. Chandra and Fisher [3] have shown the economic impact of the PRP, they have found that using the PRP reduced 3 to 20% of the total costs compared with considering each objective separately.

Thereafter, many studied have been focused on this domain to figure out the different features of the PRP: capacited PRP, single or multi item, single vehicle, homogeneous or heterogeneous fleet etc., proposing different formulations and algorithms to solve it.

Only few of exact algorithms have been proposed to solve the PRP, the first among them have been proposed by Fumero and Vercellis [4] based on the lagrangian relaxation of the problem. A similar relaxation has been proposed by Solyali and Süral [5] to solve the Inventory-Routing Problem in the case where when the customer is visited, the inventory must reach its maximal capacity. Ruokokoski et al. [6] and Bard and Nananukul [7] treated the capacited PRP, where in [6] the authors used a Branch and Cut algorithm while in [7] they developed a Branch and Price algorithm considering multi vehicle case. Absi et al. [8] introduced a heuristic algorithm to solve the capacited PRP, where they proposed an iterative approach with two phases, the first phase solve the production problem to determine the amount of product to produce and to store, the second one aims to solve a set of VRPs and TSPs to find the optimal routes for the vehicle.

Metaheuristics have always been the dominated method to solve many operational problems such as the PRP, using a research method to exploit the space and to avoid the case of

local optimum. For example, Boudia et al. [9] proposed a GRASP algorithm (Greedy Randomized Adaptive Search Procedure) with two versions: either with a reactive mechanism process or with the path relinking process. Brahimi and Tarik [10] developed a hybrid heuristic which combine a Relax and Fix method and a local search method. A similar method of Relax and fix has been implemented by Miranda et al. [11] to solve the PRP. In another context Darvish et al. [12] proposed two mathematical models: one for the IRP (inventory routing problem) and one for the PRP, one time minimizing the total costs, and the other time minimizing the carbone emissions, and finally they compared their results to show the conflict between the objectives.

Multi-objective optimization is a part of the combinatory optimization, that consists in optimizing simultaneously many objectives for the same problem, which are most of time contradictory, it has been applied to solve various kinds of problems in different domains using different methods.

Sazvar et al. [13] proposed a new replenishment policy in a centralized supply chain for deteriorating items using a bi-objective stochastic programming, minimizing inventory and transportation costs, and the greenhouse gas emissions in the same time. Amoozad-Khalili et al. [14] presented a multi-objective cell formation problem, they considered alternative process routes and machine utilization and fuzzy demand, to minimize jointly the total cell load variation and the total costs using a scatter search algorithm. Bozorgi-Amiri et al. [15] proposed a robust stochastic programming model for disaster relief logistics under uncertainty. The model aims to minimize the total related costs and in the same time to maximize the affected area's satisfaction levels using a multi-objective method which is LP-metric.

In most of papers that deals with the Production-Routing Problem, the main objective was to minimize the total costs (of production, setup, inventory and transportation) but they didn't include the energetic aspect in their works even though it represents a very important side to study. So in this paper we proposed a mathematical model that treat the energetic aspect in the PRP, once in an energy minimizing model, and then in a multi-objective model to join between the classical objective and the energetic one.

III. PROBLEM DESCRIPTION AND FORMULATION

The PRP consists to determine jointly and optimally the production and the distribution planning to minimize the total costs, in the classical PRP the production facility is responsible of producing the product and replenishing a defined set of customers, so then many decisions have to be taken: when and how much to produce? How much to store? When and how much to deliver in each period? And what are the optimal routes for the vehicle?

In this paper we treat the capacited PRP, differently than other research papers we consider a production system introducing the notion of "time", here the production capacity is limited by how much the system can produce in a certain length of period, with a single vehicle and a single item, through the three cases: minimizing costs with an energetic constraint to limit the amount of energy used while producing, minimizing energy and finally multi-objective that combines both objectives.

A. Objective 01: minimizing total costs:

Sets :

N : set of nodes (node 1 present the factory).
C : set of customers.
T : sets of periods.

Parameters :

P_t : unitary production cost.
h_i : unitary holding cost (for factory and customers)
S_t : setup cost for each period t.
CU_t : vehicle utilization cost per period t
$Dem_{i,t}$: demand of customer i in period t.
CapV : vehicle capacity.
$CapS_i$: maximum inventory holding capacity for node i.
$Stock_{i,0}$: initial inventories for factory and customers.
Duree t: length of period t.
Tprod: unitary processing time
Setup: necessary setup time for the system.
Power1: power of the system while operating.
$Power2_t$: power of the system while setting up the production in each period t.
$Maxp_t$: maximum power allowed in period t.

Decision variables

Z_t: if there is production in the factory in period t
R_t: the amount produced in period t.
$Qliv_{i,t}$: amount of item delivered to customer i in period t
$Stock_{i,t}$: inventory level in node i in period t.
$Charge_{i,j,t}$: the load of the vehicle while travelling from node i to node j in period t.
$Y_{i,t}$: if the customer i is served in period t.
$X_{i,j,t}$: if the vehicle travel from node i to node j in period t.
$Emax_t$: energy used in period t.

Mathematical model 01:

$$Min \sum_{i\in N}\sum_{t\in T} h_i * Stock_{i,t} + \sum_{t\in T} CU_t * Y_{1,t} + \sum_{t\in T} S_t * Z_t + \sum_{t\in T} P_t * R_t \quad (1).$$

Subject to:

$$Stock_{1,t} = Stock_{1,t-1} + R_t - \sum_{i\in C} Qliv_{i,t} \quad ,\forall t \in T - \{1\} \quad (2).$$

$$Stock_{i,t} = Stock_{i,t-1} + Qliv_{i,t} - Dem_{i,t} \quad ,\forall i \in C, \quad \forall t \in T - \{1\} \quad (3).$$

$$Stock_{i,t} \leq CapS_i \quad ,\forall i \in N, \quad \forall t \in T \quad (4).$$

$$Qliv_{i,t} \leq (CapS_i + Dem_{i,t}) * Y_{i,t} \quad ,\forall i \in C, \quad \forall t \in T \quad (5).$$

$$\sum_{i \in C} Qliv_{i,t} \leq CapV * Y_{1,t} \quad ,\forall t \in T \quad (6).$$

$$Charge_{i,j,t} \leq CapV * X_{i,j,t} \quad ,\forall i \in N, \quad \forall j \in N, \quad \forall t \in T \quad (7).$$

$$\sum_{j \in N} X_{i,j,t} + \sum_{j \in N} X_{j,i,t} = 2 * Y_{i,t} \quad ,\forall i \in N, \quad \forall t \in T \quad (8).$$

$$\sum_{j \in c} Charge_{1,j,t} = \sum_{j \in C} Qliv_{j,t} \quad ,\forall t \in T \quad (9).$$

$$\sum_{i \in N} Charge_{i,j,t} - \sum_{i \in N} Charge_{j,i,t} = Qliv_{j,t} \quad ,\forall j \in C, \quad \forall t \in T \quad (10).$$

$$Charge_{i,1,t} = 0 \quad ,\forall i \in C, \quad \forall t \in T \quad (11).$$

$$Charge_{i,i,t} = 0 \quad ,\forall i \in N, \quad \forall t \in T \quad (12).$$

$$X_{i,i,t} = 0 \quad ,\forall i \in N, \quad \forall t \in T \quad (13).$$

$$R_t \leq Z_t * \sum_{i \in C} \sum_{\substack{t' \in T \\ t'=t}} Dem_{i,t'} \quad ,\forall t \in T \quad (14).$$

$$\sum_{t \in T} Qliv_{i,t} = \sum_{t \in T} Dem_{i,t} \quad ,\forall t \in T \quad (15).$$

$$\sum_{i \in N} X_{i,j,t} = Y_{j,t} \quad ,\forall j \in N, \forall t \in T \quad (16).$$

$$\sum_{i \in N} \sum_{j \in C} X_{i,j,t} \geq \sum_{j \in C} Qliv_{j,t}/CapV \quad ,\forall t \in T \quad (17).$$

$$R_t * Tprod + Z_t * setup \leq Duree_t \quad ,\forall t \in T \quad (18).$$

$$Emax_t \leq Maxp_t \quad ,\forall t \in T \quad (19).$$

$$Emax_t \geq (power1 + power2_t) * Z_t \quad ,\forall t \in T \quad (20).$$

$$X_{i,j,t}, Y_{j,t}, Z_t \in \{0,1\} \quad ,\forall i \in N, \quad \forall j \in N, \quad \forall t \in T \quad (21).$$

$$Charge_{i,j,t}, Stock_{i,t}, R_t \geq 0 \quad ,\forall i \in N, \quad \forall j \in N, \quad \forall t \in T \quad (22).$$

$$Qliv_{j,t} \geq 0 \quad ,\forall j \in C, \quad \forall t \in T \quad (23).$$

The objective function (1) minimizes the total costs which include, inventory, transportation, production and setup costs, (transportation cost is calculated beside the vehicle utilization cost, every period the vehicle is used, a cost is occurred). Constraints (2) and (3) ensures the inventory balancing at both of the factory and customers. Constraints (4) and (5) ensures that the inventory capacity is not exceeded. Constraints (6) and (7) ensures that the vehicle capacity is respected. Constraint (8) indicates that when the vehicle visits a node, it must leave it. Constraints (9) and (10) calculate the load of the vehicle according to the amount delivered. Constraints (11) and (12) indicate that there is no load between the node and itself nor while coming back to the production facility. Constraint (13) insures that there's no arc between the node and itself. Constraints (14) and (15) ensures that the quantity produced and delivered respect the customer's demand. Constraint (16) is a route constraint and constraint (17) is the fractional capacity constraint for the sub-tour elimination. Constraints (18) limit the quantity produced by the length of the period. Constraints (19) and (20) calculate and limit the power used by the system. and finally constraints (21), (22) and (23) indicate the variables nature.

B. Objective 02: minimizing the energy consumed:

In this part we introduce another model that aims to minimize the energy consumed by the production system, we add to the previous model:

Decision variables

$Consom1_t$: energy consumption by the system while operating in period t

$Consom2_t$: energy consumption by the system while setting up the production in each period t.

Mathematical model 02:

$$Min \sum_{t \in T} consom1_t * R_t + \sum_{t \in T} consom2_t * Z_t \quad (24).$$

Subject to: (2) to (23), (25) and (26).

$$Consom1_t = power1 * Tprod \quad ,\forall t \in T, \quad (25).$$

$$Consom2_t = power2_t * setup \quad ,\forall t \in T \quad (26).$$

The objective function (24) replace the objective function (1) in the first model, it aims to minimize the total energy consumed by the system in the two phases: setup and production. And finally Constraints (25) and (26) calculate the energy consumed by the system in both phases production and setup respectively.

C. Objective 03: multi-objective case:

In this part we introduce another model that aims to minimize the energy consumed, these objectives are contradictory and expressed with different measure units and can't be joint directly in one function, so we need a special procedure to make it work. Here we will use LP-metric method to put the two objectives together in one function.

$$Min\ \ w1\frac{ob1-ob1^{*}}{ob1^{*}}+w2\frac{ob2-ob2^{*}}{ob2^{*}}+\ldots+wn\frac{obn-obn^{*}}{obn^{*}}$$

This equation presents the LP-metric function, ob1, ob2...obn present the objective functions to be optimized simultaneously, $ob1^{*}$, $ob2^{*}$… obn^{*}, are the objective values correspondent to their functions, and finally w1, w2..wn are relative weights to the objective functions where $0 \leq wn \leq 1$ and $\sum wn = 1$.

Then in our case the formulation of the problem will include all the parameters and variables of both of models with:

$$Ob1 = \sum_{i\in N}\sum_{t\in T} h_i * Stock_{i,t} + \sum_{t\in T} CU_t * Y_{1,t} + \sum_{t\in T} S_t * Z_t + \sum_{t\in T} P_t * R_t \quad (27).$$

$$ob2 = \sum_{t\in T} consom1_t * R_t + \sum_{t\in T} consom2_t * Z_t \quad (28).$$

Mathematical model 03:

Now the mathematical model become:

$$Min\quad w1\frac{ob1-ob1^{*}}{ob1^{*}}+\ w2\frac{ob2-ob2^{*}}{ob2^{*}} \quad (29).$$

Subject to: (2) to (23) , (25) to (28).

Where (29) is the multi-objective function, w1 and w2 are weights that make us control the importance of each function. Constraint (27) calculate the objective value of cost function, and constraint (28) calculate the objective value of energy function. So this model includes both models constraints and objective function is the LP-metric function

IV. EXPERIMENTATIONS

Now we test the present study through three different scenarios to show the efficiency of our modeling:

TABLE I: DATA IN COMMUN IN ALL SCENARIOS

		P1	P 2	P 3	P 4	P5
Demand	**C1**	100	0	200	0	0
	C2	0	300	0	150	0
	C3	50	0	250	0	100
	C4	150	100	0	0	200
	C5	0	200	100	100	0
$Maxp_t$		500	200	400	800	600
P_t		2	2	2	2	2
$Duree_t$		1000	1000	1000	1000	1000
$CapS_i$		1200	500	500	500	500
$S_{i,0}$		400	200	200	200	200
Tprod		**Setup**	**Power1**	**CapV**		
1		10	10	800		

TABLE II: DATA OF THE DIFFERENT SCENARIOS

Scenario 01						
		P1	**P2**	**P3**	**P4**	**P5**
$Power2_t$		200	260	250	300	500
S_t		50	30	25	30	50
CU_t		100	200	300	120	200
	Factory	**C1**	**C2**	**C3**	**C4**	**C5**
h_i	0	0	0	0	0	0
Weights		**W1**			**W2**	
		0.5			0.5	
Scenario 02						
$Power2_t$		150	200	500	400	500
S_t		25	15	30	55	25
CU_t		150	100	300	500	120
	Factory	**C1**	**C2**	**C3**	**C4**	**C5**
h_i	10	0	0	0	0	0
Weights		**W1**			**W2**	
		0.7			0.3	
Scenario 03						
$Power2_t$		500	100	150	400	600
S_t		10	25	35	55	25
CU_t		150	100	300	500	120
	Factory	**C1**	**C2**	**C3**	**C4**	**C5**
h_i	0	10	10	10	10	10
weights		**W1**			**W2**	
		0.3			0.7	

TABLE I presents the common data between the three scenarios including the customers demand, production cost, maximum energy allowed, period length, inventory capacity, initial inventories, processing and setup time, production power and the vehicle capacity.

TABLE II concerns the data for each scenario which include setup power, vehicle utilization cost, inventory holding cost and finally weights of the objective functions in the LP-metric model. In scenario 01 we put equal weights to both objectives to have a balance between them, in the second scenario gave more importance to the first function by putting higher weight (70%) compared to (30%) to the second objective which is considered as a weak weight, finally in the last scenario we reverse those weights to give more importance to the second objective

V. RESULTS

In this section we present the experimental results of the scenarios generated previously. These results are obtained by using the solver "Cplex", which are mentioned in the following

table that conclude all the scenarios, where in each scenario we indicated the objective value of each function, the amount produced and delivered in each period to each customer. In multi-objective study we calculate the LP-metric objective value as well as the value of both of mono-objective functions to compare with the results obtained.

TABLE III: RESULTS OF THE SCENARIOS

		Scenario 01						Scenario 02						Scenario 03					
Min cost	**Ob1**	3 695						**Ob1**	4 100					**Ob1**	28 800				
	Rt	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**	
		990	/	610	/	/			50	/	/	800	750			/	610	/	990	/	
	Qliv	300	/	/	/	/	**C1**	**Qliv**	100	/	/	/	200	**C1**	**Qliv**	/	/	100	100	100	**C1**
		250	/	/	/	200	**C2**		100	/	/	200	150	**C2**		/	100	/	150	200	**C2**
		100	/	/	300	/	**C3**		100	/	/	300	/	**C3**		/	/	100	/	300	**C3**
		50	/	/	200	200	**C4**		50	/	/	/	400	**C4**		/	50	/	400	/	**C4**
		100	/	/	300	/	**C5**		100	/	/	300	/	**C5**		/	/	100	100	200	**C5**
Min energy	**Ob2**	20 500						**Ob2**	21 500					**Ob2**	18 500				
	Rt	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**	
		610	/	990	/	/			990	/	/	610	/			/	610	990	/	/	
	Qliv	/	210	/	/	90	**C1**	**Qliv**	300	/	/	/	/	**C1**	**Qliv**	250	/	/	/	50	**C1**
		/	450	/	/	/	**C2**		150	/	/	300	/	**C2**		150	/	300	/	/	**C2**
		/	/	100	300	/	**C3**		0	350	50	/	/	**C3**		/	/	100	/	300	**C3**
		50	/	/	400	/	**C4**		50	/	/	/	400	**C4**		/	50	/	/	400	**C4**
		300	/	/	100	/	**C5**		300	/	/	/	100	**C5**		/	/	400	/	/	**C5**
LP-metric	**LP-ob**	0						**LP-ob**	0.07					**LP-ob**	0.036				
	Ob1	3 695						**Ob1**	4 100					**Ob1**	32 280				
	Ob2	20 500						**Ob1**	26 500					**Ob2**	18 500				
	Rt	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**		**Rt**	**P1**	**P2**	**P3**	**P4**	**P5**	
		990	/	610	/	/			50	/	/	800	750			/	610	990	/	/	
	Qliv	300	/	/	/	/	**C1**	**Qliv**	100	/	/	/	200	**C1**	**Qliv**	/	/	300	/	/	**C1**
		250	/	/	200	/	**C2**		100	/	/	150	200	**C2**		/	100	/	350	/	**C2**
		100	/	/	/	300	**C3**		100	/	/	300	/	**C3**		/	/	150	/	250	**C3**
		50	/	/	/	400	**C4**		50	/	/	/	400	**C4**		/	50	/	50	350	**C4**
		100	/	/	300	/	**C5**		100	/	/	300	/	**C5**		/	/	200	/	200	**C5**

VI. DISCUSSION

In this section we present the experimental results of the three scenarios including the first model of cost minimization, model of energy minimization and finally the LP-metric method determining the objectives values of the two objective considering the different weights.

In scenario 01, results of the first model (min cost) show that because the setup cost was lower in periods 1 and 3 then the production occurred in those periods, and the delivery was planned in period 1, 4 and 5 because the of the vehicle utilization cost. The second model gives almost the same results of the first model where the production occurred in period 1 and 3 (but with different amounts) where the power consumption was lower, then the distribution happened in all periods because the transportation cost is not taken into consideration in the energy minimization model. In the LP-metric case the objective value of both functions minimizing costs and energy function was exactly the same to the independent functions, that's why the production decision was in the same periods that to both models, but the amounts like the first model, while the delivery decision was in the same periods as the first model but also with different amounts to optimize the transportation cost. Here the objective functions weights were equal that's the reason why the LP-metric try to satisfy both objectives in the same time with equal proportions.

In the first objective in the second scenario, and with the same reasoning the production passes in the periods 1, 4 and 5 to respect the energetic constraints, and the delivery in the same periods of production (period 1,4 and 5) because of the inventory holding cost at the factory, there was the delivery of the initial inventory and then the amounts produced are delivered in the same periods of production. In the second objective of the energy minimization, the production took place in periods 1 and 4 that have the minimum power consumption according to the power required for setting-up the system, and the delivery is happened in all periods because the delivery decisions are not considered in this model. In the multi objective case and because the weight of the first objective was higher than the second, the solver tried to find a compromised solution that satisfy the first objective the most with a 70% rate that's why the solution was close to that of the second one with the same objective value. There was a small amount production in period 1, and with the initial inventory there was delivery to avoid the inventory holding cost at the factory and the

stock-out at the customers, because the next production was in period 4 then in period 5, where the amounts produced are delivered in the same periods.

In the first function of the last scenario, the production occurred in period 2 and 4 even though they have a big setup cost because of the energetic constraints, but because the inventory holding cost is null at the production facility and important at the customers, the delivery plan started from period 2 to period 5 because customers had initial inventories and the inventory holding cost was more important than the vehicle utilization, while in the second objective, production occurred in period 2 and 3 because of the energy consumption, and as usual the delivery happens in each period except in period 4. In the LP-metric model, we put important weight for the energy function (70%) that's why the production decision and objective value are the same as the second objective, delivery decisions are optimized according to the first model, because there was no optimization in the delivery part in the second model. So in this case, even though we give more important to objective 02 but the second objective was not ignored but taken into consideration.

Comparing the three scenarios according to the multi-objective function, in the first scenario we have put equal weights for both objectives (50% vs. 50%) so the LP-metric function tried to find the compromised solutions among them, that's why the model found a solution that satisfy both objectives with the same objective values. In the second scenario we gave more importance to the first objective (30% vs. 30%) that's why the solution was closer to that of the first model, while it was far to that of the second one (objective value of the second model was 21 500 where in the LP-metric model was 26 500). In the last scenario we favorited the second objective, so in the final results, the solution was closer to that of the second model with the same objective value of the second function, but also didn't ignore the second objective finding a solution that satisfy both objectives but the second one the most.

VII. CONCLUSION

In this paper, we studied the Production-Routing Problem which serves to jointly optimize production, inventory and distribution decisions, with a new approach that gathers the economic and the energetic side. We introduced the energetic aspect in our study to show the tradeoff between the terms: cost and energy where the energy is expressed by the power consumption in the production phase where the delivery decision is not optimized efficiently while in the classical PRP all decisions are taken basing on cost minimization where each activity is expressed by a related cost. The PRP here is treated with a different way where first the two objectives of the same problem and with the same data are optimized separately, these objectives are not expressed with the same measure unit and cannot be combined in the same objective, so we introduced a multi-objective procedure called: LP-metric.

LP-metric is one among the methods that are used to solve multi-objective methods that aims to optimize simultaneously two objectives or more for the same problem. This method is based on putting weights (importance degree) for each objective from 0 to 1, equal weights are used in the case where we aim that the final results satisfy the objectives with the same chances.

Our results show that LP-metric finds a compromise solution to satisfy all objectives even when they are contradictory, they also show that putting a "**weak**" weight for an objective doesn't mean "**ignoring**" it and that the final result is related to that of the higher weighted objective only, but related to all objectives, unless the weight is not 0, that objective is taken into consideration even with weak proportion.

VIII.REFERENCES

[1] Y. Adulyasak, J.-F. Cordeau et R. Jans, «The production routing problem: A review of formulations and solution algorithms,», 2014.

[2] P. Chandra, «A dynamic distribution model with warehouse and customer replenishment,» 1993.

[3] P. Chandra et M. L. Fisher, «Coordination of production and distribution planning,» 1994.

[4] F. Fumero et C. Vercellis, «Syncronized development of production,inventory, and distribution schedules,» 1999.

[5] O. Solyali et H. Süral, «A relaxation based solution approach for the inventory control and vehicle routing problem in vendor managed systems,» 2009.

[6] M. Ruokokoski, O. Solyali, J.-F. Cordeau, R. Jans et S. H, «Efficient formulations and a branch-and-cut algorithm for a production-Routing Problem,» 2010.

[7] J. Bard et N. Nananukul, «A branch-and-price algorithm for an integrated production and inventory routing problem,» 2010.

[8] N. Absi, C. Archetti, S. Dauzère-Pérès et D. Feillet, «A Two-Phase Iterative Heuristic Approach for the Production Routing Problem,» 2015.

[9] M. Boudia, M. A. Ould Louly et C. Prins, «A reactive grasp and path relinking for a combined production–distribution problem,» 2007.

[10] N. Brahimi et A. Tarik, «Multi-item production routing problem with backordering: a milp approach,» 2016.

[11] P. L. Miranda, R. Morabito et D. Ferreira, «Optimization model for a production, inventory, distribution and routing problem in small furniture companies,» 2018.

[12] M. Darvish, C. Archetti et L. Coelho, «Trade-offs between environmental and economic performance in production and inventory-routing problems,», 2018.

[13] Z. Sazvar, S. Mirzapour Al-e-hashem, A. Baboli et M. Akbari Jokar, «A bi-objective stochastic programming model for acentralized green supply chain with deteriorating products,» 2013.

[14] H. Amoozad-Khalili, M. Ranjbar-Bourani et S. Mirzapour Al-e-Hashem, «Fuzzy Demand Consideration in a Multi-Objective Dynamic Cell Formation Problem Using a Robust Scatter Search,» 2010.

[15] A. Bozorgi-Amiri, S. M. Jabalameli et S. M. J. . Mirzapour Al-e-Hashem, «A multi-objective robust stochastic programming model for disaster relief logistics under uncertainty,» 2011.

Multi objective simulated annealing for identical parallel machines with deterioration effect and resources consumption

SEKKAL Norelhouda
Manufacturing Engineering Laboratory of Tlemcen University of Tlemcen
Tlemcen, Algeria
Sekal.nor@hotmail.fr

BELKAID Fayçal Manufacturing Engineering Laboratory of Tlemcen
University of Tlemcen
Tlemcen, Algeria
f_belkaid@yahoo.fr

BOUFELLOUH Radhwane
Engineering Laboratory of Tlemcen
University of Tlemcen
Tlemcen, Algeria
ra_boufellouh@enst.dz

***Abstract*—Scheduling problems with a variable processing time have become a popular topic in the last decades, models with constant processing time are not appropriate for most of industrial system. In this paper we treat an identical parallel machine scheduling problem in which job processing time depend on it position in the machine and each one requires an amount of resources to be processed. This amount of resources depends on the processing time width of the job. To describe the problem more clearly, a mathematical programming model is presented. This model represents realistic situation, in which jobs affectation and resources consumption decision are considered simultaneously. But due to the complexity of the problem, the model is not able to find solution for medium and large instances. However, we adapt to the problem a metaheuristic based on multi objective simulated annealing. Moreover, we propose a specific initial solution to start the algorithm. We test the proposed algorithm through three sizes of instances. The simulations show that the algorithm performances depend on the objectives weights.**

Keywords—Parallel machine scheduling, deterioration, resources consumption, multi objective simulates-annealing.

I. INTRODUCTION (*HEADING 1*)

In a competitive word, in which production is automated in great part, the success of a factory depends on its management. And the best management is done by considering all parameters all the parameters that affect the smooth running of the production. To do that, production scheduling is a fundamental need in manufacturing industry. We mean by production scheduling assignment of jobs to machines, allocation of resources, planning of preventive maintenance and consideration of other constraints, to minimize a given objective generally relative to cost or time.

Traditional papers dealing with scheduling problems, considered jobs with constant processing time. However, in real world application, machines efficiency varies and may decrease because of usage and age. In the other hand, jobs also may deteriorate while waiting to be processed. That means that job processing times are defined by a function of their starting time or/and the position to which they are affected. Scheduling in this manner is known as the deteriorating job scheduling problem. Since it was introduced by [1] this trend has been widely adopted by researchers.

In addition, when considering a scheduling problem, machines are not the only resources to manage. However, the processing of jobs may require others resources, such as raw material, tools, energy…etc. Therefore, an optimal exploitation of these resources can reduce costs and make more profit for the system.

In most papers considering scheduling problems with resources constraints, the consumption of resources is independent on processing time, while it is not always the case. When these resources are electricity, fuel or a machine lubricant, the amount of resources consumed depends on the width of processing time.

The particular problem addressed in this study explores the area of a deteriorating job with processing time-dependent resources consumption, which distinguishes this study from most other studies. It's assumed that jobs must be processed on parallel machine, under certain conditions, to minimize both makespan and the total consumed resources cost.

For parallel machine scheduling problem, the most common objective is the minimization of the so-called *makespan. E*ach job has to be processed by only one machine, and a machine can process one job at a time, this problem has been mathematical proved to be NP-hard. In this study two function are to be minimized, makespan and cost which make the problem much more complex, where the need of adapting a meta-heuristic (multi objective simulated annealing) to solve the problem.

In this paper, a parallel machine problem with deterioration and processing time-dependent resources consumption to minimize makespan and total consumed resources cost is treated. A literature revue is made in the next section. Then mathematical model is presented in the third section. Finally, multi objective simulated annealing "MOSA" is used as resolution approach.

II. LITERATURE REVIEW

Parallel machine scheduling problem with deteriorating jobs has been widely studied in last decades. There are authors who have leaned toward a processing time functions that

Identify applicable funding agency here. If none, delete this text box.

depend on process starting time and others who have leaned toward a processing time functions that depend on the position of the job on the machine.

Authors who considered a starting time dependent processing time are:

[2] treated an identical parallel machines problem with linear jobs deterioration to minimize the total absolute differences in completion times and the total absolute differences in completion times , and presented an algorithm to solve the problem in O(n log n) time. [3] considered a parallel machine scheduling problem with start time dependent processing time to minimize makespan. [4] also considered a parallel machine scheduling problem, in which the processing time of a job is a simple linear function of its starting time. [5] proposed a fully polynomial time approximation scheme to solve a problem of scheduling parallel machine with time dependent processing time. [6] reviews models where processing time is a time dependent function. [7] considered a single-machine scheduling problems in which the processing time of a job is a function of its starting time and its resource allocation with the objective of minimizing makespan, total completion times, total waiting times, the total absolute differences in completion times or the total absolute differences in waiting times and total cost of allocated resources.

Others who considered a position dependent processing time are:

[8] treated an unrelated parallels machine problem, with at most one modifying activity (RMA) in a machine, to minimize the total completion time and total machine load. They considered the deterioration in the maintenance activity time, the later the activity begins the most it takes time. also [9] considered unrelated parallels machines problem with deterioration and RMAs. [10] considered a single machine treating different jobs with mixed deteriorating jobs. They treated the standards objectives of the maximum completion time the total completion, the total weighted completion time, the maximum lateness, the number of tardy jobs.[11] proposed a polynomial time algorithm to minimize makespan in a single machine, with jobs deterioration and modifying activities (RMA), they assumed that even after an RMA machines still deteriorated and doesn't get back their initial performances.

Older than deterioration effect resources constraints are also widely considered and in several ways;

[12] studied scheduling problems of parallel machine, where machine are dedicated and jobs require resource to be processed. [13] considered an identical parallel machine scheduling problem with consumable resource and proposed a genetic algorithm to solve the problem. [14] treated a parallel machine scheduling problem in which the processing of jobs requires a number of scarce resources and proposed an algorithm to solve small and medium instances. [15] worked on parallel machines system with additional resource, where the processing time of a job discreetly depends on the number of resource. also [16] and [17] considered both an unrelated parallel machines with resources consumption and the processing time of a job depends on the amount of resources allocated to it. Resources dependent processing time is one of the most considered effects recently. Unlike this and deterioration, few papers considered processing time dependent resources consumption; we mentioned among those [18] and more recently [19]

As for the resolution approach, we mentioned [20] who developed a meta-heuristic algorithm (tabu search) to solve a parallel machine bi-objective problem in a deteriorating system (total tardiness and machines deteriorating cost are the two function to minimize). [21] adapted a simulates annealing to an unrelated parallel machines problem with sequence-dependent setup times

III. PROBLEM STATEMENT

The considered system is in form of m=1…M identical parallel machines that have p=1…P positions to treat j=1…N jobs. And it possesses the following characteristics: machine treat one job at a time and each job are processed in only one position of a machine. Processing time of a job depend on it position in the machine. Jobs consume resources depending on their processing time. Resources are available in sufficient quantity at time zero.

Processing time of a job:

$$Pr_{jp} = a_j + \alpha.(p-1) \quad (1)$$

Where:

a_j is the normal processing time of the job h

α: is the position compression rate.

Consumption resources of a job:

$$r_j = \gamma.a_j + \gamma'.(p-1) \quad (2)$$

Where: γ and γ'are time compression rate. And p_j is the position of the job j.

Objective function:

The objective in this study is to minimize makespan and resources costs. The two objectives are not constrained, in others words; minimizing only makespan doesn't increase resources costs, and minimizing resources costs no longer doesn't increase makespan. On the contrary resources consumption depends on the processing time, so it is relative to minimize a time entity. This entity is minimized only if deterioration is well managed.

System parameters:

N total number of jobs

M total number of machines

a_j : normal processing time of job j

α : deteriorating rate of job j

γ and γ' are normal time compression rate, and added time compression rate, respectively.

Cu: unit cost of resources R.

Pr_{pk}: Processing time of the job in position p machine k.

r_{pk}: Amount of resources consumed in position p, machine k.

Cr: Total consumed resources cost.

C_k : completion time.

C_{max} : makespan

Decision variable:

$$X_{jpk} = \begin{cases} 1 & if\ j\ is\ processed\ in\ position\ p\ machine\ m, \\ 0 & otherwise, \end{cases}$$

Mathematical model:

$$Min\ F1 = C_{max} \quad (3)$$

$$Min\ F2 = Cr \quad (4)$$

$$\sum_{j=1}^{N} X_{jpk} \leq 1 \quad p=1\ldots P,\ k=1\ldots M \quad (5)$$

$$\sum_{k=1}^{M}\sum_{p=1}^{N} X_{jpk} = 1 \quad j=1\ldots N \quad (6)$$

$$X_{jpk} \leq \sum_{j=1}^{N} X_{j(p-1)k} \quad p=1\ldots P,\ k=1\ldots M \quad (7)$$

$$Pr_{pk} = \sum_{j=1}^{N} X_{jpk}\,a_j + \alpha_j(p-1) \quad p=1\ldots P,\ k=1\ldots M \quad (8)$$

$$r_{pk} = \gamma \sum_{j=1}^{N} X_{jpk}\,a_j + \gamma'\left(\alpha_j(p-1)\right) \quad p=1\ldots P,\ k=1\ldots M \quad (9)$$

$$Cr = \sum_{k=1}^{M}\sum_{p=1}^{N} r_{pk} * Cu \quad (10)$$

$$C_k = \sum_{p=1}^{P} Pr_{pk} \quad k=1\ldots M \quad (11)$$

$$Cmax \geq C_k \quad (12)$$

Equations (3) and (4) are objective functions; makespan and resources costs. Equation (5), (6) and (7) ensures that one job at a time. Equations (6) and (7) ensure that jobs are treated each one in one position. Equation (8) defines the actual processing time that depend on the normal processing time of the job, it position in the machine and it compression rate. Equation (9) defines the amount of consumed resources, also depending on normal processing time of the job and the time added by deterioration effect. Equations (10) and (11) define the total comsumed resources cost and completion time, respectively. Finally, equation (12) defines the makespan.

IV. RESOLUTION APPROACH

It is agreed that parallel machines problems are NP-hard. Adding deterioration and resources consumption to a parallel machine problem, we conclude that our system is NP-hard, and even solver only gives solutions for small instances. For these reasons a meta-heuristic (simulated annealing) is adopted as a resolution approach to our system.

Simulated Annealing SA belongs to meta-heuristics based on solution modification. It is characterized by using one solution by iteration which reduce the number of parameters to fixe and make it more simple to adapt to our system.

Inspired by the physical process of metals cooling, this algorithm has been established in 1980 by [22]. The cooling process aims to order the atoms in most regular structure. For this, cooling rate has crucial impact on final structure. So the SA's algorithm proceeds.

The algorithm :

An initial structure is defined (randomly or using a heuristic), also initial temperature is defined and must be high enough to allow enough flexibility to the algorithm, but not so high as to slow down the algorithm. The movement from a structure to another is done if the quality of the new structure is judged better than the quality of the initial one:

- If the new structure gives better results than the initial one.
- Else, if the new structure probability is higher than acceptance probability.

The acceptance probability, otherwise named *metropolis acceptance criterion*, is considered to escape the local optimum. The structure probability at a given temperature is calculated according to the following distribution:

$$P = e^{-\frac{\Delta S}{T_n}} \quad (13)$$

Where: $\Delta S = S_n - S_i$

And T_n is the temperature in the nth iteration.

Simulated Annealing is a finite number of iteration repeated under decreasing temperature according to the cooling ratio α.

It is noted that:

- Initial solution is perturbated to obtain new solution
- A structure is a solution to the problem.
- During each iteration temperature is decreased using the following relation $T_n = \alpha . T_{n-1}$
- Iterations stop with the stopping criterion is achieved.

SA algorithm is an algorithm for mono-objective optimization problem, to adapt it to a multi-objective optimization problem [23] proposes (under the name of MOSA multi objective simulated annealing) an algorithm that differs from SA only in the method for calculating the acceptance probability. This method is as follow:

First, probability is calculated for each function k

$$P_i(k) = \begin{cases} e^{-\frac{\Delta S}{T}} \ if\ \Delta S > 0 \\ 1 \quad else \end{cases} \quad (14)$$

Then, after all probabilities are defined, they are aggregated either by choosing the smallest of them or by calculating their weighted product:

$$P = \prod_{k=1}^{K} P_i(k)^{w_k} \quad (15)$$

$$P = min\{P_i(k)\} \quad (16)$$

Where w_k is the weight of the function k.

Algorithm setting :

Before applying MOSA algorithm, it is important to well put the initial parameters (initial structure and initial temperature), and well define the cooling ratio, acceptance probability and stopping criterion.

- Initial structure or initial solution: The initial solution may be chosen randomly as it can be defined using a heuristic (LPT, SPT…etc). Using a heuristic or a method to define the initial solution increases the

probability of finding the best solution reducing runtime of the algorithm.

In this study, the initial solution is defined so as at least one objective function is minimized. The simplest function to minimize is total consumed resources cost. This can be done by processing in each machine the same number of jobs. So defining the initial solution consists of dividing jobs on machines.

- Initial temperature T_i , cooling ratio α, acceptance probability P_a and stopping criterion δ: Cooling ratio is fixed to 0.99 enough near from 1 to sweep a sufficient solution in a small temperature range. Initial temperature and acceptance probability are fixed to 8000 and 0.60, respectively. Stopping criterion is theoretically achieved when T=0, but in practice that may not be achieved, so a very small positive number δ is considered, so that when T_n =δ algorithm stop repetition.

The fundamental steps of the algorithm:

- Coding:

While α is common to all jobs, the sequence of jobs in a machine doesn't make a difference for either functions. So in the first line of the chromosome, jobs are positioned from first to last without any order. In the second line the machines where every job is processed are defined. In the third line the positions where every job is put are defined. See Table I

TABLE I. EXAMPLE OF A CHROMOSOME USING FOR CODING SOLUTION

Jobs	J1	J2	J3	J4	J5	…….	Jn
Machines	2	m-1	m-1	1	M	…….	2
positions	1	1	2	1	1	……	2

- Neighborhood generation:

To generate a neighborhood, the classical method is used "mutation". A mutation is a random change made on a case (one or several cases) selected randomly. From the chromosome in tab.1, there are three lines that accept mutation in their cases. But the only affectation that imports for the two objectives functions is the affectation of jobs to machines, second line.

The first and third line define the order and the position of jobs on machine, otherwise said they define the sequence of jobs on a machine, which doesn't import for any of the two functions. So only one random mutation is made on a case of the second line.

- Correction and evaluation of the solution:

To avoid any empty position in machines, after every mutation the positions of all machines are verified and redefined. The value of makespan and total consumed resources cost are calculated for the new solution, and if it doesn't give better results than the initial solution also probability is calculated, according to acceptance probability the solution is judged to be accepted or not

Algorithm:

INITIALIZE ($S_i, T_i, \alpha, P_a, \delta$)

While ($T_n < \delta$):

i- Neighborhood generation:
S_n=mutation (S_i)

ii- Fitness:
Calculate OF (S_n), and ΔS= OF (S_n) - OF (S_i)

iii- Select Si for the next iteration:

- If (ΔS<0) Then
$$S_i = S_n$$
- Else:

{ Calculate the probability
$$P = e^{-\frac{\Delta S}{T_n}}$$
- If (P_a>=P) then:
{S_i=S_n accepted}
- Else { Si=Si rejected}

iv- Update (T_i)

Endwhile

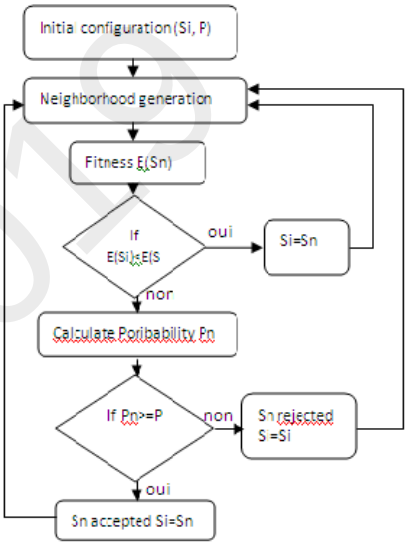


Fig. 1. Simulated annealing UML

Numerical analysis of the algorithm performances:

In this section results of several computational experiments are shown, with small, medium and large instances. Processing time are generated randomly according to a normal distribution p_j=U(5,50). Rates α, γ and γ' are common to all jobs, and generated according to a distribution U(0,1), U(0,1) and U(0,1.5), respectively. Resources' unit cost is generated randomly according to U(1,5). The weight of the objective functions are considered (w_1, w_2)=(0.3, 0.7); (0.5, 0.5); (0.7, 0.3) to show the impact of the weight of both functions.

The following tables II, III and IV resume results of simulations made with MOSA on small, medium and large instances, respectively. For each instance and each value of (w_1, w_2), the average of 10 simulations and the best value are considered.

In table II instances are from a small size, results given by MOSA are compared with results given by a solver (Lingo 10).

For instances of medium and large sizes the solver cannot find a solution. So in tab.III and IV, the average is compared with min and max (best and worst results given by MOSA). Two deviations are considered,

TABLE II. RESULTS OBTAINED FROM MOSA AND LINGO FOR SMALL INSTANCES

W1w2	0.30.7				0.50.5				0.70.3			
	MOSA		Lingo	Dev.	MOSA		Lingo	Dev.	MOSA		Lingo	Dev.
	Aver.	Min			Aver.	min			Aver.	min		
1	91.1	90.6	90.6	0.58	95.0	94.2	94.2	0.87	99	97.9	97.9	1.12
2	177	175.8	175.8	0.72	162.8	161.7	161.7	0.71	150.26	147.7	147.7	1.73
3	295.2	294.8	294.8	0.15	245.3	244.1	244.1	0.50	194.4	193.5	293.4	0.48
4	193.7	192.7	192.7	0.56	181.4	180.8	180.8	0.34	170	169	169	0.69
5	372.6	371.7	371.7	0.25	290.5	289.4	289.4	0.38	208.9	207.2	207.2	0.80

TABLE III. RESULTS OBTAINED FROM MOSA FOR MEDIUM INSTANCE

W1w2	0.30.7					0.50.5					0.70.3				
	Aver.	Min	Max	D1%	D2%	Aver.	Min	Max	D1%	D2%	Aver.	Min	Max	D1%	D2%
1	1216.5	1209	1229.9	0.59	2.2	947.5	908.3	937	1.02	3.15	615.1	606.5	634.4	1.4	4.6
2	993.7	989.6	1003.29	0.42	1.38	749.4	744.4	759.8	0.68	2.0	505.4	497.4	532.4	1.6	7
3	406.1	401.9	414.2	1	3	343.4	332.1	348	2.9	4.7	273.5	262.2	285.2	4.3	8.7
4	1126.9	1124.3	1136.3	0.23	1	847.7	841.8	858	0.7	1.9	563.8	559.3	574.7	0.8	2.7
5	663.5	660.4	669.4	0.47	1.36	524.8	517.1	541	1.4	4.6	378.4	373.9	410	1.2	9.6

TABLE IV. RESULTS OBTAINED FROM MOSA FOR LARGE INSTANCES

W1w2	0.30.7					0.50.5					0.70.3				
	Aver.	Min	Max	D1%	D2%	Aver.	Min	Max	D1%	D2%	Aver.	Min	Max	D1%	D2%
1	2767.7	2743.5	2791	0.8	1.7	2067.8	2035.4	2110	1.5	3.6	1366.1	1327.2	1413	2.9	6.4
2	5204.8	5175.6	5250	0.5	1.4	3827.6	3786.6	3869	1	2.1	2428.4	2397.6	2486	1.2	3.6
3	1545.1	1677.6	1722	1.1	2.6	1311.3	1285.6	1364	2	6	939.39	893.6	994.7	5.1	11.3
4	3882.6	3874.2	3910	0.2	0.9	2877.9	2840.4	2922	1.3	2.8	1858.9	1806.6	2022	2.8	11.9
5	3474.9	3438.8	3562	1	3.5	2572.8	2540.5	2647	1.2	4.1	1671.3	1642.3	1722	1.7	4.8

D1 between average and min and D2 between max and min

- In table II It's noticed that the optimal value of the objective function are given by MOSA at least once in ten simulations. Also the deviation between average and the optimal value is generally less then 1%. It is also noticed that the most w_1 increase and w_2 decrease, the deviation between average and optimum increases.

In table.III, medium instances results are resumed, it is noticed that the deviation between best solution and the average D1 doesn't exceed 1% when $w_2 > w_1$, this deviation increases as the value of w_1 increases and w_2 decreases. As for D2 (deviation between min and max), for the three cases, it is more important than D1, but it remains always lower for a greater value of w_2.

For table.IV, the same observations are made as for table III. But it's added that the deviation is more important for large instances than for medium instances. To illustrate this note, in Figure I, the average deviation between average and best obtained solution is calculated for small, medium and large instances, and it is shown that the average deviation increases slightly from small to medium instances and also from medium to large instances.

For the three cases of small, medium and large instances, it's denoted that for $w_2 > w_1$ the algorithm gives best results. When the total consumed resources cost has more weight in the objective function any deviation in the makespan doesn't have a great influence on the final results. Conversely to this, when makespan weighs more than total consumed resources cost in the objective function the slightest deviation makes a difference. As mentioned before, the initial solution is put so that total consumed resources cost is minimized (by assigning the same number of jobs to each machine) without bearing interest to makespan, and this is the reason why the deviation is the smallest when $w_2 > w_1$.

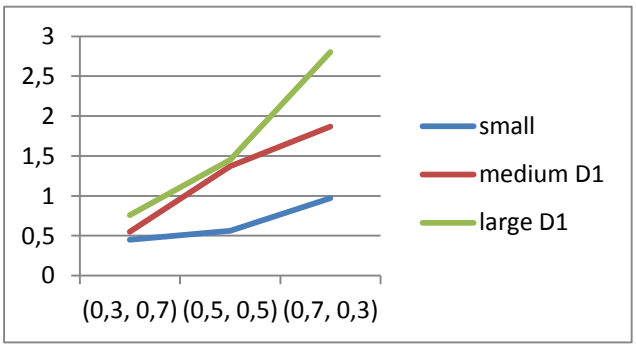


Fig. 2. Average deviation of small, medium and large instances results

V. CONCLUSION:

In this study, we presented a mathematical model of parallel machine system with deterioration and processing time

dependent resources consumption. Due to the complexity of the system the mathematical model cannot handle efficiently the problem in reasonable time. For that, we adapted a multi objective meta-heuristic algorithm (MOSA) to the system to solve the problem of jobs' assigning to minimize resources cost and makespan..

Through a numerical analysis, we showed that, for small instances algorithm find the optimal solution at least once in ten simulations. We don't have a solution to take as reference for medium and large instances, but we can conclude from results given by small instances that the algorithm converges to the optimum even for more important size of instances. Despite this we can conclude that results given by the algorithm for medium and large instances are satisfactory view that deviation still acceptable.

Using MOSA algorithm as a resolution approach allows us to weight the two objective functions relative to each others, and so to understand the behavior of the system according to the importance of each function.

The study of parallel machines scheduling problem with unconventional constraint remains a fundamental axis in the management and organization of factories, and continue to attract the interest of a considerable number of researchers. Therefore, some effort should be made to improve the study in term of:

- System's consideration:

In this study we assumed that machines are always functional and resources are available in unlimited quantities. However, in reality machines are subject to break down and require maintenance. Also resources are not always available and may arrive at a specifics moments. Thus, there is a possibility to add breakdowns constraints and maintenance to the system or/and modify the mathematical model so that the limit amount of resources is respected.

- Resolution approach :

In this study we adapt an multi objective simulated annealing with a specific initial solution, and we noticed that the algorithm loses his performance when makespan weighs more than resources cost. So more research need to be done to find the adequate approach for better results

REFERENCES

[1] Browne, S., Yechiali, U., 1989. SCHEDULING DETERIORATING JOBS ON A SINGLE PROCESSOR.

[2] Huang, X., Wang, M.-Z., 2011. Parallel identical machines scheduling with deteriorating jobs and total absolute differences penalties. Applied Mathematical Modelling 35, 1349–1353. https://doi.org/10.1016/j.apm.2010.09.013

[3] Ouazene, Y., Yalaoui, F., 2017. Identical parallel machine scheduling with time-dependent processing times. Theoretical Computer Science. https://doi.org/10.1016/j.tcs.2017.12.001

[4] Chen, Z.-L., 1996. Parallel machine scheduling with time dependent processing times. Discrete Applied Mathematics 70, 81–93. https://doi.org/10.1016/0166-218X(96)00102-3

[5] Ji, M., Cheng, T.C.E., 2008. Parallel-machine scheduling with simple linear deterioration to minimize total completion time. European Journal of Operational Research 188, 342–347. https://doi.org/10.1016/j.ejor.2007.04.050

[6] Alidaee, B., Womer, N., 1999. Scheduling with time dependent processing times: Review and extensions. Journal of the Operational Research Society.

[7] Wei, C.-M., Wang, J.-B., Ji, P., 2012. Single-machine scheduling with time-and-resource-dependent processing times. Applied Mathematical Modelling 36, 792–798. https://doi.org/10.1016/j.apm.2011.07.005

[8] Cheng, T.C.E., Hsu, C.-J., Yang, D.-L., 2011. Unrelated parallel-machine scheduling with deteriorating maintenance activities. Computers & Industrial Engineering 60, 602–605. https://doi.org/10.1016/j.cie.2010.12.017

[9] Gara-Ali, A., Finke, G., Espinouse, M.-L., 2016. Parallel-machine scheduling with maintenance: Praising the assignment problem. European Journal of Operational Research 252, 90–97. https://doi.org/10.1016/j.ejor.2015.12.047

[10] Gawiejnowicz, S., Lin, B.M.T., 2010. Scheduling time-dependent jobs under mixed deterioration. Applied Mathematics and Computation 216, 438–447. https://doi.org/10.1016/j.amc.2010.01.037

[11] Rustogi, K., Strusevich, V.A., 2012. Single machine scheduling with general positional deterioration and rate-modifying maintenance. Omega 40, 791–804. https://doi.org/10.1016/j.omega.2011.12.007

[12] Kellerer, H., Strusevich, V.A., 2003. Scheduling problems for parallel dedicated machines under multiple resource constraints. Discrete Applied Mathematics 133, 45–68. https://doi.org/10.1016/S0166-218X(03)00433-5

[13] Belkaid, F., Yalaoui, F., Sari, Z., 2016. An Efficient Approach for the Reentrant Parallel Machines Scheduling Problem under Consumable Resources Constraints: International Journal of Information Systems and Supply Chain Management 9, 1–25. https://doi.org/10.4018/IJISSCM.2016070101

[14] Fanjul-Peyro, L., Perea, F., Ruiz, R., 2017. Models and matheuristics for the unrelated parallel machine scheduling problem with additional resources. European Journal of Operational Research 260, 482–493. https://doi.org/10.1016/j.ejor.2017.01.002

[15] Daniels Richard L., Barbara J., H., Joseph B., M., 1996. Scheduling Parallel Manufacturing Cells with Resource Flexibility. Management science 42.

[16] Edis, E.B., Oguz, C., 2012. Parallel machine scheduling with flexible resources. Computers & Industrial Engineering 63, 433–447. https://doi.org/10.1016/j.cie.2012.03.018

[17] Hsieh, P.-H., Yang, S.-J., Yang, D.-L., 2015. Decision support for unrelated parallel machine scheduling with discrete controllable processing times. Applied Soft Computing 30, 475–483. https://doi.org/10.1016/j.asoc.2015.01.028

[18] Vickson, R.G., 1980. Two Single Machine Sequencing Problems Involving Controllable Job Processing Times. A I I E Transactions 12, 258–262. https://doi.org/10.1080/05695558008974515

[19] Tigane, M., Dahane, M., Boudhar, M., 2018. Multiobjective approach for deteriorating jobs scheduling for a sustainable manufacturing system. The International Journal of Advanced Manufacturing Technology. https://doi.org/10.1007/s00170-018-3043-1

[20] Mazdeh, M.M., Zaerpour, F., Zareei, A., Hajinezhad, A., 2010. Parallel machines scheduling to minimize job tardiness and machine deteriorating cost with deteriorating jobs. Applied Mathematical Modelling 34, 1498–1510. https://doi.org/10.1016/j.apm.2009.08.023

[21] Kim, D.-W., Kim, K.-H., Jang, W., Chen, F.F., 2002. Unrelated parallel machine scheduling with setup times using simulated annealing. Robotic and computer integrated manufacturing 223–231.

[22] S. Kirkpatrick, C. D. Gelatt Jr, M. P. Vecchi, 1983. Optimization by Simulated Annealing. science 671–680.

[23] Ulungu, E.L., Teghem, J., Fortemps, P.H., Tuyttens, D., 1999. MOSA method: a tool for solving multiobjective combinatorial optimization problems. Journal of Multi-Criteria Decision Analysis 8, 221–236. https://doi.org/10.1002/(SICI)1099-1360(199907)8:4<221::AID-MCDA247>3.0.CO;2-O

Visual decision support for Breast Cancer Recurrence

BENHACINE Fatima Zohra
Laboratoire d'Informatique Oran (LIO)
Département d'informatique
Université Oran 1 Ahmed Ben Bella
Oran, Algerie
benhacine.fatima@gmail.com

ATMANI Baghdad
Laboratoire d'Informatique Oran (LIO)
Département d'informatique
Université Oran 1 Ahmed Ben Bella
Oran, Algerie
baghdad.atmani@gmail.com

ABDELOUHAB Fawzia Zohra
Laboratoire d'Informatique Oran (LIO)
Département d'informatique
Université Oran 1 Ahmed Ben Bella
Oran, Algerie
fzabdelouhab@gmail.com

***Abstract*— Our study focuses mainly on visual decision support, particularly on the visualization of association rules for the detection of the non-recurrence of cancer. Data mining aims to extract as much relevant information as possible from a large amount of data. It is done automatically, or by exploring the data using interactive visualization tools. The article presents our Visual4AR approach, which jointly uses the Cellular Automaton CASI (Cellular Automaton for Symbolic Induction) and 2D colored matrices for the detection of non-recurrence of cancer.**

The interest is double. On the one hand, this provides an accessible visual representation. On the other hand, it allows to infer and interact with the system in order to easily select the most relevant rules from the mass of rules.

Keywords*— *Association Rules, Data Mining, visualization, Interactive visualization system, Cellular Automaton, Colored 2D Matrix, Boolean Modeling, Breast Cancer Recurrence

I. INTRODUCTION

Data Mining (DM) is a process of extracting knowledge from a very large amount of data. The principle is to search structures linking these data. This search can be done automatically, for example by using algorithms whose purpose is to find association rules such as Apriori [1]. An advantage of this algorithmic approach is its exhaustiveness, thanks to which all the association rules, which satisfy constraints on a set of metrics, will be found [2]. However, the number of rules built can sometimes be larger than the initial amount of data. In this case, we are faced with another problem of data mining, which consists in identifying the most relevant subsets of rules. To solve the problem of rule mass, visualization is presented as a potential solution in the post-processing of knowledge models. Indeed, the visualization of information helps human beings to acquire and increase their knowledge and to guide their reasoning through their perspective capacities.

In this article, we propose to use an interactive visualization tool Visual4AR (Visual for Association Rules), [3], for medical decision support including the detection of non-recurrence of cancer.

Visual4AR [3] is an automatic approach to data mining where knowledge extraction is done first under Boolean modeling using the CASI cellular Automaton [4] and then visual exploration by data selection.

The first part of this article deals with data mining visualization based on work related to visualization analysis and visual perception. Then, we detail our approach and show how to use our Visual4AR tool [3] to produce a Boolean visualization of association rules to explore and filter relevant rules in the breast cancer field.

II. RELATED WORK

The clinical decision support system (CDSS), are playing increasingly important roles in medical practice by helping physicians or other medical professionals making clinical decisions. CDSS are having a greater influence about the care process. They are expected to improve the medical care quality; their impact should intensify due to increasing capacity for more efficient data processing [5]

Breast Cancer is a complex disease characterized by multiple variables obtained from several data-sources, such as clinical, genetic or image sources. Over the past decades, various studies have tried to predict the outcome of breast cancer with the support of these data, and big advances have been done in this direction [6].

[7] presents a semantic web approach to develop a clinical decision support system to support family physicians to provide breast cancer follow-up care. their approach involved the computerization and execution of a breast cancer follow-up clinical practice guideline. The computerization of the clinical practice guideline led to the development of a breast cancer ontology. Their breast cancer ontology which models the knowledge inherent within the breast cancer follow-up clinical practice guideline - the breast cancer ontology serves as the knowledge source to determine patient specific recommendations.

[8] describes a personalized clinical decision support system for cancer care. Data from evidence-based medical knowledge services are semantically linked with electronic health records and presented to consultants at the point-of-care.

[9] develop the Decision Support System for Making Personalized Assessments and Recommendations Concerning Breast Cancer Patients (DPAC), which is a CDSS learned from data that recommends the optimal treatment decisions based on a patient's features.

As we have pointed out in previous work, most visualization methods are not suitable for representing large

sets of patterns. They become unusable when the number of patterns to be displayed is too large and few displays give an overview of the pattern sets. Finally, no method is suitable to really explain how to deduce certain rules or the presence of other rules [3].

We are interested in our work on the Two Dimensional Matrix (2D) for the simplicity to represent the rules on two colors, blue for antecedent and red for consequent. Nonetheless, in the presence of a significant number of association rules this representation also becomes illegible, and the rules overlap. Occultation problems of such representation become inevitable. To overcome the problem of occultation on the one hand and to simplify the internal representation of manipulated knowledge as well as to improve reasoning and control on the other hand, we have opted for rule optimization using boolean modeling offered by the Cellular Automaton (CASI), and its inference engine to explain the reasoning of some deductions.

In our approach we based on some aspects of the CASI Cellular Automaton: its simplicity to express knowledge in rules and facts, its efficiency in optimizing storage space and execution time. The latter is a particular model of dynamic and discrete systems able to acquire, represent and process extracted knowledge in Boolean form [10].

The originality of our system is essentially in the combination between the 2D matrix and the CASI Cellular Automaton to prove its effectiveness in a new field which is visualization.

III. BREAST CANCER AND VISUALIZATION

Recurrent breast cancer is a cancer that often recurs 2 to 15 years after initial treatment to eliminate all cancer cells [11]. Breast cancer is a major public health issue in most developed countries. Over the past decade, significant progress has been made in its prevention and management. The continuation of this dynamic requires regular monitoring of epidemiological data, both to organize care chains and to guide the implementation of new actions and assess their effectiveness.

A. Data preparation

We used the real breast cancer medical data set [12], provided by the UCI repository, whose data set has a total of 286 bodies representing women with breast cancer described by 10 attributes (TABLE I).

The objective of this study is to help medical experts detect breast cancer recurrence by providing solid rules from the patient database.

TABLE I. ATTRIBUTE DESCRIPTION

Attributes	Descriptions	Value
age	patient's age at diagnosis	10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70-79, 80-89, 90-99
Menopause	the patient is pre or post-menopausal at the time of diagnosis;	lt40, ge40, premeno
Tumor size	the size of the largest tumour diameter (in mm) of the excised tumour	0-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44,45-49, 50-54,55-59
Inv-nodes	the number (range 0-39) of axillary lymph nodes that contain metastatic breast cancer	0-2, 3-5, 6-8, 9-11, 12-14, 15-17, 18-20, 21-23, 24-26,27-29, 30-32,33-35, 36-39
Node caps	if the cancer is metastasized to a lymph node outside the tumor's original site, it may remain "contained" by the lymph node capsule from which the tumor can replace the lymph node	yes, no
Degree of malignancy	the histological grade (range 1-3) of the tumor. Tumors that are grade 1 are composed of cells, while neoplastic, retain many of their usual characteristics. While grade 3 tumors are composed of cells that are highly abnormal;	1, 2, 3
Breast	breast cancer can definitely occur in the other breast;	left, right
Breast quadrant	the breast can be divided into four quadrants, using the nipple as a central point;	left-up, left-low, right-up, right-low, central
irradiation	Radiotherapy is a treatment that uses high-energy x-rays to destroy cancer cells;	yes, no
Class	without recurrence or recurrence based on the recurrence of breast cancer symptoms in patients after treatment.	no-recurrence-events, recurrence-events

B. Extraction of association rules

One of the important applications of data mining is to extract association rules from a large amount of data. In this study, we use the Apriori algorithm [1]. An association rule links different attributes describing the data.

A rule A → B means: if the attributes contained in A are present in a given transaction, then the attributes contained in B are also present. Apriori [1] only takes into account nominal attributes. Two basic metrics are used to evaluate and retain interesting patterns:

- The support: probability of the simultaneous occurrence of A and B
- Confidence: probability that B is true if A is true

Confidence measures the probability that the rule is true. The support measures the probability that a rule has to be applied. These measurements serve as constraints for the Apriori algorithm. Given the thresholds for support and trust,

Apriori guarantees the extraction of all association rules whose levels of trust and support are above the thresholds.

The user intervenes at two distinct points in the process. First, it sets the support and confidence thresholds that will be used during the research. Once completed, it is the sole judge of the results that are presented, the direct consequence is that some items are not taken into account if they are not well represented in the database. To remedy this, the user can provide feedback by modifying the initial parameters and lowering the support, which shows the flexibility of our system.

By applying the Apriori algorithm [1] we have generated a set of association rules, an extract of which is given in the following Fig.1.

R1: inv-nodes=0-2 deg-malig=1 Class=no-recurrence-events-->node-caps=no
R2: inv-nodes=0-2 deg-malig=1 Class=no-recurrence-events -->irradiat=no
R3:inv-nodes=0-2 deg-malig=1 Class=no-recurrence-events-->node-caps=no irradiat=no
R4: deg-malig=1 irradiat=no Class=no-recurrence-events -->node-caps=no
R5:deg-malig=1 irradiat=no Class=no-recurrence-events -->inv-nodes=0-2
R6:deg-malig=1 irradiat=no Class=no-recurrence-events-->inv-nodes=0-2 node-caps=no
R7:inv-nodes=0-2 breast-quad=left_low irradiat=no Class=no-recurrence-events-->node-caps=no
R8: age=50-59 inv-nodes=0-2 node-caps=no Class=no-recurrence-events-->irradiat=no
R9:inv-nodes=0-2 deg-malig=1 irradiat=no --> Class=no-recurrence-events node-caps=no
R10:menopause=ge40 inv-nodes=0-2 irradiat=no -->node-caps=no

Fig. 1. Extract from the Association Rules

This extract of rules (Fig. 1) will give rise to a knowledge base (Fig. 2) on which the CASI cellular machine will operate. This knowledge base is made up of a Facts base which are all proposals such as "inv-nodes=0-2" or "Class=no-recurrence-events" and a rule base which are all the extracted rules.

A	inv-nodes=0-2
B	deg-malig=1
C	Class=no-recurrence-events
D	node-caps=no
E	irradiat=no
F	breast-quad=left_low
G	age=50-59
H	menopause=ge40

R_1	A, B →C
R_2	F,D → A
R_3	D,E→B
R_4	B,D→F
R_5	E,F → D
R_6	E,F →B
R_7	B,F →G

Fig. 2. The knowledge base

C. Boolean rule modeling

The Boolean modeling, we used, which respects the principle of cellular automata, is described by four Boolean matrices and two transition functions that simulate the operation of an inference engine [3]. The initial state of the machine is given by the CelFact matrix expressing the fact base and the CelRule matrix expressing the rule base.

Fait	EF	IF	SF
A	0	1	0
B	0	1	0
C	0	1	0
D	0	1	0
E	0	1	0
F	0	1	0
G	0	1	0
CELFACT			

Règles	ER	IR	SR
R_1	0	1	1
R_2	0	1	1
R_3	0	1	1
R_4	0	1	1
R_5	0	1	1
R_6	0	1	1
R_7	0	1	1
CELRULE			

Fig. 3. Initial state of CelFact and CelRules

In the Knowledge Base in Fig. 2, the root node of the first rule represents the initial fact. The initial configuration of the machine for the initial knowledge base is given by the initial state of CelFact and CelRule (Fig. 3) and RE and RS (Fig. 4).

RE	R_1	R_2	R_3	R_4	R_5	R_6	R_7
A	1	0	0	0	0	0	0
B	1	0	0	1	0	0	1
C	0	0	0	1	0	0	0
D	0	1	1	0	0	0	0
E	0	0	1	0	1	1	0
F	0	1	0	0	1	1	1
G	0	0	0	0	0	0	0

RS	R_1	R_2	R_3	R_4	R_5	R_6	R_7
A	0	1	0	0	0	0	0
B	0	0	0	0	0	1	0
C	1	0	1	0	0	0	0
D	0	0	0	0	1	0	0
E	0	0	0	1	0	0	0
F	0	0	0	0	0	0	0
G	0	0	0	0	0	0	1

Fig. 4. RE and RS initial status

D. Colored 2D matrix generated by CASI

As we explained [3] we used the visualization by the colored 2D matrix. The latter is given in the form of "rules-to-itemset" matrices, where each row corresponds to an item and each column to a rule.

E. Boolean visualization

The CASI cellular Automaton through its inference engine will allow us to refine all these rules by eliminating at each iteration the irrelevant rules according to the user's needs. By choosing the visualization by selection, the user validates his starting item; the rule to validate. The change of state of the machine, which behaves like a cellular automaton at this level, is done by transition functions whose role is to simulate the operation of a front linkage.

In this case, we have defined two functions, δ_fact and δ_rules which operate on the CelFact and CelRule matrices respectively.

- The Transition Function δ_{fact} is given as follows:

$$(EF, IF, SF, ER, IR, SR) \xrightarrow{\delta_{fact}} (EF, IF, EF, ER + (R_E{}^t . EF), IR, SR)$$

- The Transition Function δ_{rules} is given as follows

$$(EF, IF, SF, ER, IR, SR) \xrightarrow{\delta_{rules}} (EF + (RS . ER), IF, SF, ER, IR, \text{^}ER)$$

Where the matrix RE^T refers to the transposition of RE and ^ER denotes the negation of the Boolean vector ER.

The 2D matrix displays the applicable rules step by step, we note that the rules applicable in this step are R3, we obtain a 1st visualization as follows (Fig. 5):

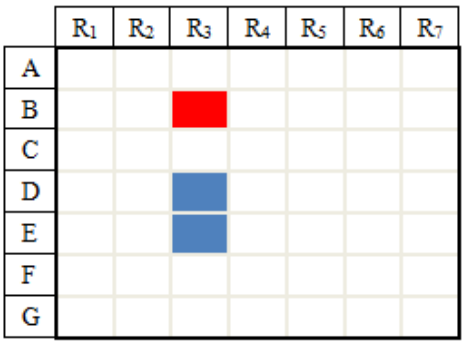


Fig. 5. Visualization of the R3 (relevant rule)

The validation of R3 led to the validation of R4 (Fig. 6) :

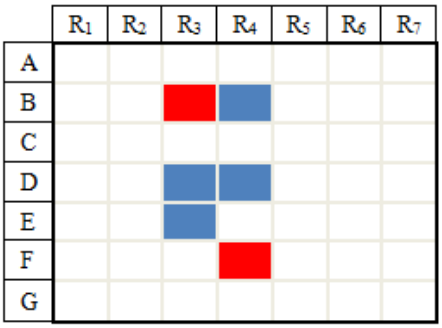


Fig. 6. Visualization of R3 and R4 rules

The process repeats from one configuration to another until there are no more candidate rules (including ER=0) to select. The rules will be deactivated as their conclusions are validated (Fig. 7).

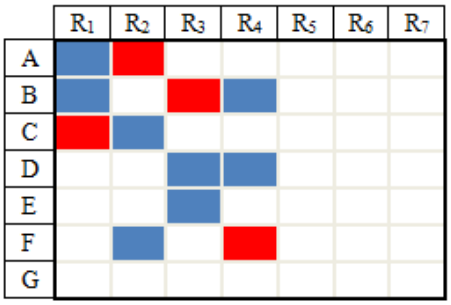


Fig. 7. Visualization of R3 and R4 rules

Suppose that $G = \{G_0, G_1, G_2 \dots \dots G_q\}$ is the set of configurations of our cellular automaton. The discrete evolution of the PLC, from one generation to another, is defined by G_0 , $G_1, .. G_q$, $ou\ G_{i+1} = \Delta(G_i)$ from every graph of the automaton we generate the colored 2D matrix displaying, thus, the relevant rules of this iteration [3].

The user has the choice to stop the system if he considers that he has reached his goal, otherwise he continues his exploration towards the other graphs. This interactivity with the system allows the user, on the one hand, to easily select the most relevant rules from the mass of rules and, on the other hand, he is integrated into the knowledge search loop.

IV. CASE STUDY

In the previous section we detailed our approach on a reduced sample of the breast cancer base [11]. For this part we will show the behavior of Visual4AR on this basis. To do this, we first performed several tests over several series of experiments to choose the best model containing an adequate number of association rules (TABLE II).

TABLE II. ASSOCIATION RULE SETS

Test	Min Support	Confidence Min	number of rules
Series1	10%	95%	257
Series 2	10%	90%	300
Series 3	10%	85%	805
Series 4	10%	80%	1177
Series 5	9%	90%	648
Series 6	9%	80%	1436
Series 7	9%	70%	2167
Series 8	8%	90%	809
Series 9	8%	80%	1768
Series 10	8%	70%	2673
Series 11	10%	50%	2954

For experiment N°1 we used the 2nd series (TABLE II) with 300 interesting rules (Fig. 8) that we cannot all interpret, to retrieve the rules that can meet our objective which is the detection of non-recurrence of cancer using the Boolean matrix and 2D colored visualization to solve the occultation problem.

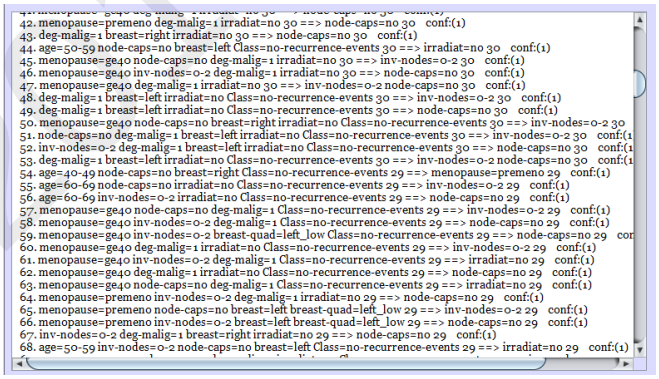


Fig. 8. Visualization of R3 and R4 rules

The rules triggered are R200, R206, R207, R208, R209, R219, R220, R220, R220. (Fig. 9). These rules were presented to a doctor who helped us interpret them.

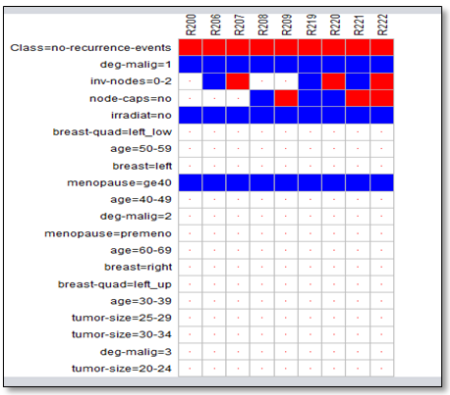


Fig. 9. Colored 2D visualization

We consulted with doctors from the University Hospital Centre of Oran City and helped us to interpret the relevant rules as follows:

Rule 200 is interpreted as follows: menopause=ge40 deg-malig=1 irradiat=no→ Class=no-recurrence-events, conf:(0.97) informs us that the degree of grade 1 malignancy, post-menopausal, the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the non-recurrence of cancer (presented by a red color).

Rule 206: menopause=ge40 inv-nodes=0-2 deg-malig=1 irradiat=no → Class=no-recurrence-events, conf:(0.97) informs us that the post-menopausal and lymph node range between 0-2, the degree of grade 1 malignancy and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the non-recurrence of cancer (presented by a red color).

Rule 207: menopause=ge40 deg-malig=1 irradiat=no → inv-nodes=0-2 Class=no-recurrence-events conf:(0.97) informs us that post-menopausal, grade 1 malignancy and lack of radiotherapy (presented by a blue color) can cause cancer in 97% of cases and that the lymph node range is between 0-2, (presented by a red color).

Rule 208: menopause=ge40 node-caps=no deg-malig=1 irradiat=no → Class=no-recurrence-events conf:(0.97) informs us that post-menopausal and the absence of a tumor that has replaced all lymph nodes with a degree of malignancy of grade 1 and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases of no-recurrence of cancer (presented by a red color).

Rule 209: menopause=ge40 deg-malig=1 irradiat=no → node-caps=no→ Class=no-recurrence-events (0,97) informs us that post-menopausal and grade 1 malignancy and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the non-recurrence of cancer and the absence of a tumor that has replaced all lymph nodes (presented by a red color).

Rule 219: menopause=ge40 inv-nodes=0-2 node-caps=no deg-malig=1 irradiat=no → Class=no-recurrence-events conf: (0.97) shows that the postmenopausal and lymph node range between 0-2 the absence of a tumor that has replaced all lymph nodes, the degree of grade 1 malignancy, and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the non-recurrence of cancer (presented by a red color).

Rule 220: menopause=ge40 node-caps=no deg-malig=1 irradiat=no→ inv-nodes=0-2 Class=no-recurrence-events conf: (0.97); indicates that post-menopausal and the absence of a tumor that has replaced all lymph nodes, the degree of grade 1 malignancy, and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases of non-recurrence of cancer with a lymph node range between 0-2 (presented by a red color).

Rule 221: menopause=ge40 inv-nodes=0-2 deg-malig=1 irradiat=no → node-caps=no Class=no-recurrence-events conf: (0.97); shows that the postmenopausal and lymph node range between 0-2, the degree of grade 1 malignancy, and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the non-recurrence of cancer with an absence of a tumor that has replaced all the lymph nodes (presented by a red color).

Rule 222: menopause=ge40 deg-malig=1 irradiat=no → inv-nodes=0-2 node-caps=no Class=no-recurrence-events; conf (0,97) shows that post-menopausal, grade 1 malignancy, and the absence of radiotherapy (presented by a blue color) can cause in 97% of cases the no-recurrence of cancer with an absence of a tumor that has replaced all lymph nodes and the lymph node range between 0-2 (presented by a red color).

Experiment N°2 makes it possible to detect cancer recurrence, we used series 11 (TABLE II) with a set of 2954 rules, the initial fact base is (recurrence-event). We launched a back chaining to get explanations, we obtained 2 relevant rules R2450 and R2673 (Fig. 10).

Fig. 10. second Experience

Rule 2450: node-caps=yes → Class=recurrence-events conf: (0.554). This rule shows that the absence of radiotherapy (blue colour) is responsible for 55% of cases of recurrence of cancer presented by a red color).

Rule 2637:deg-malig=3 → Class=recurrence-events conf: (0.53). This rule shows that the degree of grade 3 malignancy is responsible for 53% of cases of cancer recurrence (presented by a red color).

V. Conclusion

In this paper, we are interested in accessible visual representation and in inference and interaction with the system in order to easily select the most relevant rules from the mass of rules. We have highlighted the combined use of visualization techniques and Data Mining in a CDSS. We used an approach coupling a colored 2D matrix with a Boolean modeling to visualize association rules for the detection of the non-recurrence of cancer. However, some improvements can be considered such as establishing grayscales to colors according to trust and rule support. Submit to experts, sets of rules from medical fields to validate our approach. And finally, we propose to evaluate our approach on other fields of application and with other visualization techniques.

REFERENCES

[1] R. Agrawal, R. Srikant. " Fast algorithms for mining association rules". In : Proc. 20th int. conf. very large data bases, VLDB. 1994. p. 487-499.

[2] G. Bothorel. Algorithmes automatiques pour la fouille visuelle de données et la visualisation de règles d'association : application aux données aéronautiques. 2014. Thèse de doctorat.

[3] F.Z Benhacine. B. Atmani. And F.Z Abdelouhab. "Contribution to the Association Rules Visualization for Decision Support: A Combined Use Between Boolean Modeling and the Colored 2D Matrix ". International Journal of Interactive Multimedia and Artificial Intelligence(inPress), 2018.

[4] B. Atmani, , B. Beldjilali, "Knowledge Discovery in Database : Induction Graph and Cellular Automaton ", Computing and Informatics Journal, 26(2), 171-197, 2007.

[5] G. Leroy, H. Chen. "Introduction to the special issue on decision support in medicine", Decision Support Systems, 43 (4), 1203-1206, 2007.

[6] N. Larburu, M. Arrue, N. Muro. "Exploring Breast Cancer Patterns for Different Outcomes using Artificial Intelligence.'' In : 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom). IEEE, 2018. p. 1-6.

[7] S.R Abidi. "Ontology-based modeling of breast cancer follow-up clinical practice guideline for providing clinical decision support. " In : Twentieth IEEE International Symposium on Computer-Based Medical Systems (CBMS'07). IEEE, 2007. p. 542-547.

[8] X. Jiang, A. Wells, A. Brufsky. "A clinical decision support system learned from data to personalize treatment recommendations towards preventing breast cancer metastasis". *PloS one*, 2019, vol. 14, no 3, p. e0213292.

[9] H, Bernhard G. et P Walsh. Personalised clinical decision support for cancer care. In : *Semantic Applications*. Springer Vieweg, Berlin, Heidelberg, 2018. p. 125-143.

[10] F.Z Abdelouhab. "Conception et réalisation d'un système cellulaire d'alimentation d'entrepôt de données à partir des sources de données hétérogènes". 2017. Thèse de doctorat. Université Oran 1 Ahmed Ben Bella.

[11] H. Ruijuan. "Medical data mining based on association rules". Computer and Information Science, 2010, vol. 3, no 4, p. 104.

[12] Breast Cancer, http://tunedit.org/repo/UCI/breast-cancer.arff (2013).

Trust and Context Aware Splitting Approach for Improving Prediction in Recommender System

El yebdri Zeyneb
AbouBekr Belkaid University of Tlemcen
LRIT Laborartory
Tlemcen, Algeria
zeyneb.elyebdri @mail.univ-tlemcen.dz

Benslimane Sidi Mohamed
Ecole Superieure en Informatique
LabRI Laborartory
Sidi Belabes, Algeria
s.benslimane@esi-sba.dz

Lahfa Fedoua
AbouBekr Belkaid University of Tlemcen
LRIT Laborartory
Tlemcen, Algeria
lahfa.f @mail.univ-tlemcen.dz

***Abstract*— Context aware splitting approach (CASA) is one of the most efficient pre-filtering approach of context aware recommender system (CARS). The idea under this approach is that an item (and/or user) evaluated in two different contextual conditions is considered as two different items (or/and users). Subsequently, after this process, we can use any traditional algorithm of recommender system. However, this approach suffers from data sparsity after the splitting process. In this paper, we propose to add trust information, in regard to Trust-Aware recommender systems (TARS), as another valuable information source to overcome this problem. Our solution is exploiting the benefits of TARS as well CASA for improving of prediction in recommender systems. Experimental results on the real-world dataset indicated satisfactory results.**

Keywords*— *recommender system, context-aware, splitting approach, trust

I. INTRODUCTION

Recommender systems (RS) provide a promising solution to the problem of information overload which has become an essential research area. RSs are generally classified into two categories: content-based systems (CB) and collaborative filtering-based systems (CF) [1]. Content-based systems are based on analyzing a set of descriptions of items previously rated by the user. Afterwards, a profile of that user's interests is created based on the features of these items [2]. Collaborative filtering systems do not rely on the content descriptions of items to predict the utility of items for a particular user, it based on the items previously rated by other users. CF plays an important role in RS, since it has been one of the most successful methods of recommendation compared to CB. Two major classes of CF algorithms can be identified: memory-based CF and model-based CF. Memory-based CF explore the user-item rating matrix in order to provide recommendations; applies neighborhood-based recommendation algorithms to predict a user's ratings depending on the ratings given by like-minded users. In contrast, Model-based CF uses the user-item ratings to learn a model, which is then used to generate online prediction [3]. The standard formulation of the recommendation problem begins with a two-dimensional (2D) matrix of ratings, organized by user and item: Users x Item → Ratings. Although these systems achieve in general a good performance in terms of prediction accuracy, but in certain domains, users' ratings are not specified in which contextual conditions the item was evaluated. Furthermore, the exact evaluation of an item can be influenced when these contextual conditions changes [1][4][5]. Hence the need to incorporate context information into account. For example, in Tourism domain, the evaluation can be high for a hotel due to a visit in summer, which is not the case for the winter. Context-Aware recommender systems (CARSs) extend the traditional formulation of the recommendation problem by incorporating context information about user-item interactions. Incorporating contexts requires that we estimate user preferences using a multidimensional rating function – R: Users x Items x Contexts →Ratings [6]. Three main approaches are proposed to incorporate the contextual information in the recommendation [1]: contextual pre-filtering, contextual post-filtering and contextual modeling. The contextual pre-filtering consist to filter relevant data according to the current context, then use any traditional recommendation algorithm on these filtered data. In contextual post-filtering, apply a recommendation algorithm to the original preference data, and after, the filter is applied to adjust the generated recommendations according to the current context. In contextual modeling, incorporate contextual information together with the user and item data for generating recommendations.

In this paper, we are interested particularly to Context Aware Splitting approach (CASA) [7] that belongs to the contextual pre-filtering approach. The context splitting algorithms are considered the most efficient and popular among the other algorithms for context-aware recommendation [8][9][10] which aim to produce a two-dimensional (2D) rating matrix. Three types of algorithms have been proposed in CASA: item splitting [8], user splitting [11] and UI splitting [9]. Item splitting algorithm considers that an item evaluated in two different contexts, so that there exist significant difference (using statistical tests called: impurity criteria [10]) are considered as two new virtual items. Consequently, we can use any traditional recommendation algorithms, which work on 2D rating matrix. Similarly, user splitting is an extension of item splitting whereas the difference is to split user into two new virtual users instead item. Likewise, UI (User-Item) splitting fuses item and user splitting. However, despite the success of CASA, this approach suffers from data sparsity problem after splitting. Data sparsity and cold start issues are the weak point of collaborative filtering because small portion of items are rated by users. As a result, the prediction becomes difficult because reliable similar users cannot be found. To overcome this limitation, several approaches have been proposed in the literature. One of the most important approaches is to incorporate trust.

In the reality, people trust the opinions of their loved ones and friends than strangers, which have some impacts on the user's choice to the target item. For these reasons, trust-based recommender systems have been proposed [12][13][14], which make more accurate recommendations based on the ratings of trusted friends.

For this, we propose to exploit both trust and context information for making better collaborative recommendations and propose a novel model that combines two approaches: context-aware splitting approach and trust–aware approach.

The reminder of the paper is organized as follows: Section 2 gives an overview of related work. The Section 3 present formalization of our proposition. Section 4 details the proposed approach. Then, our experimental results are summarized in Section 5. Finally, in Section 6, we conclude with a summary of our contributions and the future perspectives.

II. RELATED WORK

We focus our paper on three types of works: first, those who integrate social and trust network information and trust between users in traditional recommender system, secondly, those that integrate context and trust information and using contextual pre-filtering approach of CARS and thirdly propositions using context-splitting approach.

A. Trust-Aware recommender system (TARS)

Recommender system uses social network information to mitigate the data sparsity and cold start issues, which improves the accuracy of recommender system. Recent research reports have shown that employing social factors such as trust statements in RS could lead to improve the quality of recommendations [13, 14]. Two types of the trust statement are considered in trust-based recommendation methods: implicit and explicit trust statements. Explicit trust statements refer to those relationships that are directly made by users, while implicit ones are estimated by using propagation algorithm to estimate trust value. The explicit trust provided is either not available or is so sparse. Consequently, researchers suggested different ways to calculate implicit trust [16]. Jamali et al. [13] combine explicit and implicit trust but ignore similarity between users. Furthermore, in [14] they have used matrix factorization with trust propagation for recommendation. In Sinha et al [12], the authors demonstrated that, given a choice between recommendations from trusted friends and those from recommender systems, in terms of quality and usefulness, trusted friends' recommendations are preferred, even though the recommendations given by the recommender systems have a high novelty factor. Azadjalal, M. M. et al. [16] propose a novel method, which use Pareto dominance and confidence concepts to improve the accuracy of TARS and eliminate trust users that do not provide greater information, which, can be removed from the target user's trust network. Moradi, P et al. [17] enhance TARS by providing a novel method to improve prediction accuracy based on a novel reliability measure. This measure does not use the user-item rating matrix to calculate the quality of the predicted rate, but considered the trust statements. Parvin, H et al. [18] proposed a trust-based recommender algorithm called: Trust-based Collaborative Filtering using ACO (Ant Colony Optimization). In the later, ACO is used for rating prediction during user weighting. For user selection, the proposition in [18] inspired from Moradi, P et al., to compute similarity values between the target user and the others uses. In addition, Yadav, S. et al., [19] use swarm intelligence-based on meta heuristic algorithm to optimize the weights of trust metrics after analyze different trust metrics. Another model is proposed in [20], which use neural network models based on Denoising Autoencoder to improve accuracy of TARS, reduce the data sparsity, and mitigate the cold start problem, while both explicit and implicit trusts are taken into account.

B. Trust and context aware Recommender System

In order to recommend an item, [21] extend context-aware matrix factorization (CAMF) [22] which belong to contextual modeling approach to TCMF: Trust-based Context-aware Matrix Factorization for Collaborative Filtering. TCMF use both context and trust information into the baseline predictors (user bias and item bias) and user-item-context-trust interaction. This work has been extended in [23], which takes in consideration the influence of dynamic trust and the transitivity of trust on ratings simultaneously based on the social network analysis. Moreover, the work in [24] use trust and context and random walk to select more relevant items. In [25], authors combine user context, trust network of friendship, and collaborative filtering algorithm to propose a novel service recommendation for MSN (mobile social network). Furthermore, Otebolaku, A et al. [26], exploit data obtained from IoT (Internet of Things) objects to improve the quality of context-aware personalized recommendations by incorporating also trust to enforce the reliability of information sources (context and users).

C. Context-aware Splitting Approach

Context-aware splitting approach (CASA) is a collaborative filtering pre-filtering Context-aware Recommender System introduced by Baltrunas et al [10]. It is the most efficient and popular context pre-filtering techniques [9]. We distinguish three types of Splitting: Item Splitting initiated by Baltrunas et al. [8], User Splitting by Said et al. [11], User and Item (UI) Splitting [9] by Zheng et al. In CARS, user rate item in a specific contextual conditions. Contextual condition is the value that a contextual dimension can take. For example of movies data rating: Weekend and Weekday are two contextual conditions for Time contextual dimension, Cinema / Home for Location dimension. Item splitting iterates over all contextual conditions of each context dimensions and evaluates the splits based on the impurity criteria [10]. Impurity criteria aim to detect statistically significant differences among ratings to identify if contextual dimensions affects the user choice or the opinion of the user or not. Several impurity criteria are used, based on Baltrunas and Ricci's research's [8]: t_{mean}, t_{chi}, t_{prop}, and t_{IG} also Compos and al. [27] proposed another impurity criterion based on the Fisher's exact test. One splitting process is done, the original multi-dimensional rating matrix is transformed to a two-dimensional matrix. The data of this matrix is contextualized. So, context is used as filter to preselect rating profiles, and then apply the recommendation algorithms only with profiles that contains ratings in matching

contexts. As already mentioned, User splitting based on the same principle of item splitting except that split users instead of items [11]. In addition, generally, just one contextual condition is used to apply splitting, called: simple splitting; It is also possible to perform a complex split using multiple conditions across multiple context dimensions. However, there are significant costs of sparsity and potential overfitting when using multiple conditions [10]. Authors in [28] used discrete Binary Particle Swarm Optimization (BPSO) algorithm for discrete optimization purposes in the search for the optimal contextual condition combination and determined the number of contextual conditions to select for splitting the user. The work in [29], evaluate the utility of the contextual emotion factor in context-aware recommender system. One of the types of context-aware algorithms used for evaluating performance is context-aware splitting. [30] extend work of [29] and integrate more than single context (emotion) for improving recommendation. First step is to find the most appropriate contextual combination then context-aware splitting approach is used and specifically UI splitting for evaluating the effectiveness of contextual combination.

Fig. 1 provide the summary of these new approaches and each of them proved more effective than the traditional 2D recommender systems (no context and no trust information). on one hand, the Context-aware splitting approaches (CASA) improves the prediction accuracy of contextual recommendations, but nevertheless after splitting process, the original multi-dimensional rating matrix is transformed to a two-dimensional matrix which suffer from sparsity. On other hand, trust-aware approach have benefit to reduce cold start and sparsity problems. However, no previous works evaluate these algorithms with trust-aware recommendation models.

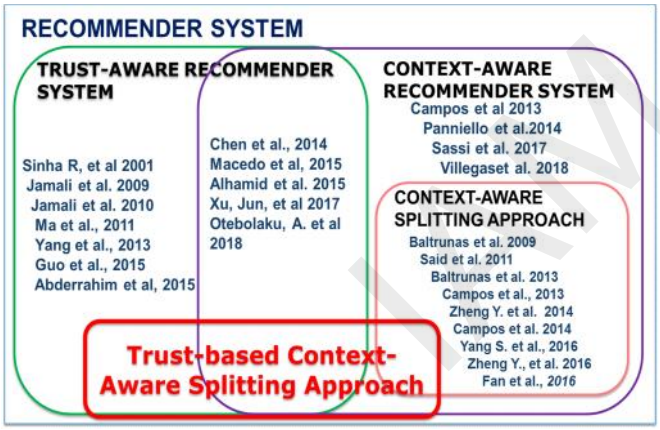


Fig. 1. A synthesis of related work

In this article, we aim to extract the benefits of both approaches to improve the quality of the recommendation by integrating two information: trust and context information. More specifically, we enhance item-splitting algorithm by integrating trust statements between users in step of prediction.

III. FORMALISATION

The basic data source of the technology in recommender system is 2D matrix. This matrix contain all ratings. These ratings are stored in a m×n matrix called rating matrix where m is the number of users, and n number of items. We denote U the set of users where $U = \{u_1, u_2, \ldots, u_m\}$, and I denote the set of items where $I = \{i_1, i_2, \ldots, i_n\}$. If a user u rates an item i, it generate a rating defined as follow: r_{ui}. When context is incorporated, recommender system extend this technology by incorporating user's contextual information as follow: U × I × C_{sit} →R, where C_{sit} represent a contextual situation, which describe the context in which the user rated the item.

Formally, Context is represented as set of contextual dimensions as follow: $C = \{C_1, C_2, \ldots, C_k\}$, where C_d present one dimension of context, such as Time. A specific value in a contextual dimension, represent contextual condition and defined as follow: $C_d = \{c_1, c_2, \ldots c_k, \ldots, c_s\}$, where s is the number of variant value of dimension C_d and c_k is one of the values that a dimension can take.

For example, in the dimension Time, there are several values (such as morning, noon, afternoon, and evening) which presented as: $C_{time} = \{morning, noon, afternoon, evening\}$. Now, the data in dataset is represented in a multi-dimensional rating space. Each rating is represented $r_{uic_{sit}}$ where C_{sit} is defined as : $\{c_1, c_2, \ldots, c_k\}$.

As example: (u_1, i_1, 5, weekend, home, friend) means u_1 rate 5 i_1 in the contextual situation: time is weekend, location is home and companion is friend.

IV. TRUST-BASED CONTEXT-AWARE SPLITTING APPROACH (TBCASA)

In this section, we describe our approach of recommender system. Fig. 2 present the overall architecture of our proposition. Our approach considers several factors for providing personalized recommendation. At the first, we include the enrichment of semantic description of items and users by collecting data from Linked Open Data (LOD). Then, we incorporate contextual information and trust network among users to predict rating of user to item. In order to have a contextualized recommendation, our approach is founded on context-aware splitting algorithms.

In the rest of paper, we interest just to item splitting algorithm.

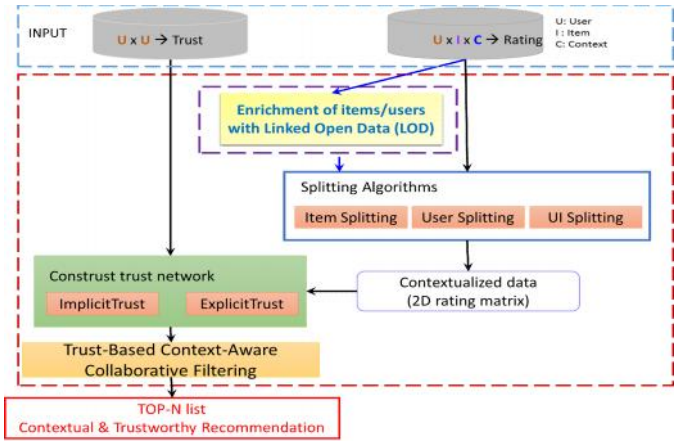


Fig. 2. Framework of our approach

Once the splitting process is completed, we focuses on generating predictions by combining trust information with a user-based technique.

A. Using Linked Open Data-Based Semantic

The objective of this step is to enhance the description of items and enrich the users' profiles. For this, we propose to use Linked Open Data (LOD), which is considered as a rich and reliable source of content information [31]. LOD plays an important role to overcome problems of cold start (new-user, new-item) and to mitigate sparsity problems of the collaborative recommender systems.

B. Selecting Items

Once we have enriched the item descriptions, and user profiles, we are interested only in the ratings that match the current context. To obtain contextualized data, we split item by employing simple item splitting, which means that use only single contextual condition for splitting because using more than one condition increase data sparsity . For this, for each item i∈I, we iterate each contextual condition c∈Cj in each contextual dimension Cj∈C. If the item i have significant differences (using impurity criteria) in the rating matrix, we create two new artificial items, then we split vector of his rating $r_i = \{r_{1i}, \dots, r_{mi}\}$ into two vectors r_{ic} and $r_{i\bar{c}}$. $r_{ic} = \{r_{1i_c}, \dots, r_{mi_c}\}$ where $r_{ui_c} = r_{ui}$ if cj=c and $r_{i\bar{c}} = \{r_{1i_{\bar{c}}}, \dots, r_{mi_{\bar{c}}}\}$ where $r_{ui_{\bar{c}}} = r_{ui}$ if cj ≠ c. For first results, we use t_{mean} as impurity criteria. We obtain as output of this step only contextualized data and two-dimensional rating matrix. However, as previously stated, the resulting rating matrix suffers from sparsity and cold start problems. To tackle these problems, we integrate, in the next step, the social trust relationships between users to compute predicted ratings

C. Construction of trust network

The objective of this step aims to incorporate social trust information into context-aware splitting approach. We include explicit and implicit trust derivable from a trust network. Explicit trust refers to the trust statements directly specified by users. For example, users in Epinions[1] and CiaoDVDs[2] can add other users into their trust lists. By contrast, implicit trust is the relationship that is not directly specified by users and that is often inferred through other information, such as propagate trust in the network by creating new relationships between users. For this, we construct trust network based on the explicit trust information available in the adjacency matrix and the Pearson correlation coefficient measure [32]. This network is directed and weighted, where, the nodes represent the users and edge represent available trust statement between the two nodes. Since the majority of users do not declare their trusted users, we propagate trust in the network by creating new relationships between users. It is reasonable to assume that, if user u trusts user b and b trusts v, u can trust v to some degree. Equation 1 is used to calculate the trust statement between pair of users [33].

$$t_{u,v} = \left(\frac{d_{max} - d_{u,v} + 1}{d_{max}}\right) \quad (1)$$

where $d_{u,v}$ is the shortest trust propagation distance between thruster u and trustee v and d_{max} is the maximum propagation distance, which represent any positive integer value. In this work, we restrict $d_{max} \leq 3$.

Therefore, the explicit trust information is generally very sparse, which not be able to predict ratings if user does not have any trusted friend for its unseen items. However, integrating also implicit trust information between users is still a challenge. For this, we infer the implicit trust relationship by employing similarity measure to assess the relationship between users. For that, we use Pearson correlation coefficient to compute similarity between two users as follows:

$$Sim_{u,v} = \frac{\sum_{i \in I_{u,v}} (\ddot{r}_{u,i} - \bar{\ddot{r}}_u).(\ddot{r}_{v,i} - \bar{\ddot{r}}_v)}{\sqrt{\sum_{i \in I_{u,v}} (\ddot{r}_{u,i} - \bar{\ddot{r}}_u)^2} \sqrt{\sum_{i \in I_{u,v}} (\ddot{r}_{v,i} - \bar{\ddot{r}}_v)^2}} \quad (2)$$

Where $I_{u,v}$ is the common set of items noted by users u and v, and $\bar{r}_u$ ($\bar{r}_v$) is the average of the rating giving by the user u (v), and $r_{u,i}$ is the note given by user u to item i. We select only Top-K nearest neighbors of the active user. The final weight $w_{u,v}$ between users *u* and *v* is calculated as follows:

$$w_{u,v} = \begin{cases} \frac{2 \times Sim_{u,v} \times t_{u,v}}{Sim_{u,v} - t_{u,v}} & if\ Sim_{u,v} \neq 0\ and\ t_{u,v} \neq 0 \\ t_{u,v} & if\ Sim_{u,v} = 0\ and\ t_{u,v} \neq 0 \\ Sim_{u,v} & if\ Sim_{u,v} \neq 0\ and\ t_{u,v} = 0 \\ 0 & if\ Sim_{u,v} = 0\ and\ t_{u,v} = 0 \end{cases} \quad (3)$$

D. Trust-based context_splitting collaborative filtering

The trust network and splitting process will be used to predict rating of the active user *u* for the item *i* . For this, we must identify the target contextual condition of the prediction, (a value c in some Cj) where cj=tc (targetcontext). The prediction is computed for the item ic if tc = c by using (4), and others if tc≠ c:

$$\breve{r}_{u,ic} = \bar{r}_u + \frac{\sum_{v \in T_u} w_{u,v} (r_{v,ic} - \bar{r}_v)}{\sum_{v \in T_u} t_{u,v}} \quad (4)$$

Where $\breve{r}_{u,i}$ denotes initial rating assigned, T_u denotes the set of trusts users of the user u, $r_{v,i}$ denotes the rating given by the user v to the item i , $w_{u,v}$ denotes weight value between user u and j and $\bar{r}_u$ ($\bar{r}_v$) is the average of the rating giving by the user u (v).

E. Recommendation

Furthermore, in the Top-N recommendations step, the algorithm selects the top-N items as recommendation list to suggest to the active user.

V. EXPERIMENTATION

This section shows experiments for analyzing the efficacy of our proposed framework. Experiments are performed on a

[1] http://www.trustlet.org/extended_epinions.html
[2] https://www.librec.net/datasets.html

real dataset. Accordingly, we present the results and compare our method with state-of-the-art recommendation methods.

A. Dataset

In this paper, CiaoDVDs is used in the experiments. In this data set, users rate various items between 1 and 5 at different timestamps. Moreover, these users can also express their trust statements with the other users. The values of the trust statements in this dataset are 0 or 1. The dataset contains 17,615 users, 16,121 movies and 72,521 ratings. In addition, there are 40,133 trust relationships between 4658 of the users.

We have enriched this dataset with different contextual concepts including day type (Weekdays, Weekends), seasons (Autumn, Winter, Spring, Summer) to evaluate our model.

B. Evaluation metric

We used a cross-validation method compatible with the above conditions. We thus performed 10-fold cross validation in all the experiments. Subsequently, we computed the accuracy of the evaluated approaches in terms of the error on rating prediction, by using Mean Absolute Error (MAE), which has widely been used in recommendation research.

$$MAE = \frac{\sum_{u,i} |r_{u,i} - p_{u,i}|}{N} \quad (5)$$

where $r_{u,i}$ is the actual rating and $p_{u,i}$ is the predicted rating. N denotes the number of tested ratings. The smaller the value of RMSE is, the more precisely the recommendation is.

C. Experimental setting

a) Splitting setting: As already mentioned, we use impurity criteria to know the best split, we use in our case tmeans [10] which estimates the statistical significance of the difference in the means of ratings in every alternative contextual condition, by using a t-test.*:*

$$T_{mean} = \left| \frac{\mu_{i_c} - \mu_{i_{\bar{c}}}}{\sqrt{\frac{{\sigma_{i_c}}^2}{n_{i_c}} + \frac{{\sigma_{i_{\bar{c}}}}^2}{n_{i_{\bar{c}}}}}} \right| \quad (7)$$

where μ_{i_c} is the average rating value of item i_c, ${\sigma_{i_c}}^2$ is the rating value variance of item $i_c$and n_{i_c} is the number of ratings given to item i_c .

Fig. 3 shows the average rating value computed over the different contexts in used dataset.

D. Experimental Results

In this section, a number of experiments are performed to evaluate the performance of the proposed method. CBCASA is compared with several well-known and state-of- the-art methods including *User-based Collaborative Filtering* (UBCF),

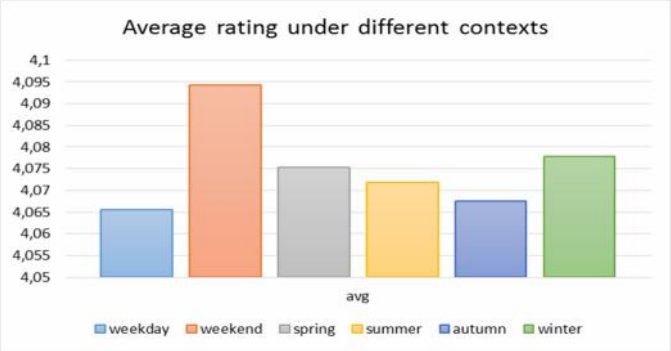


Fig. 3. Average rating under different contexts

Trust Aware collaborative filtering (TACF) which not considers contexts when measuring the similarity and trust between users, *item Splitting* (User Splitting and UI splitting for future work).

Table 1. Summarize our experiments and show the performance of the proposed TBCASA.

TABLE I. RESULT OF COMPARISON

Algorithm	UBCF	IKNN	TACF	Item splitting + UBCF	TBCASA
MAE	0,848	0,851	0,833	0,818	0,812

We observe that TBCASA exceeds UBCF that not considers trust and context information. In addition, item Splitting achieved better results compared to UBCF, while using UBCF after splitting items, which indicate that context information improve the prediction but trust information is not incorporated.

Fig. 4 represents the number of items splitted for each fold (10-fold cross validation). We see through the graph that approximately 2.64% of the items are splitted, confirming that the ratings of users are influenced by contextual information.

Fig. 4. Number of items splitted in each fold

The results also reveal that TACF outperform UBCF when considers only the influence of trust information, ignoring the influence of contextual information. Our proposed approach, which consider both trust and context information improve relevant recommendations. This improvement is explained by the addition of probability that the item is noted by user's neighbors. In addition, the opinion of these other users are qualified trustworthy because it is based on the trust statement.

VI. CONCLUSION

In this paper, we propose a novel rating prediction method that suggests items to users by using trust network among users and context of users and items. To process contextual information, we chose the splitting algorithm as one of the most effective methods for a contextualized recommendation. However, these methods do not solve the problem of data sparsity and cold start. For this, we have exploited the advantages of Trust-aware Recommender System to overcome these problems. A network of trust will be created of a target user based on trust statements between the users. Experimental results on the real-world dataset indicated satisfactory results.

As future work, we will test it on another richer dataset such as epinion. We will use not only trust information among users was used but also distrust information can be used to accurate the proposed approach.

REFERENCES

[1] Adomavicius, Gediminas, et al. "Incorporating contextual information in recommender systems using a multidimensional approach." ACM Transactions on Information Systems (TOIS) 23(1), 103-145, 2005.

[2] Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. IEEE Trans. Knowl. Data Eng. 17(6), 734–749, 2005.

[3] Yu, K., Schwaighofer, A., Tresp, V., et al.: Probabilistic memory-based collaborative filtering. IEEE Trans Knowl Data Eng 16(1), 56–69, 2004.

[4] G. Adomavicius, R. Sankaranarayanan, S. Sen, and A. Tuzhilin. Incorporating contextual information in recommender systems using a multidimensional approach. ACM Transactins on Information Systems, 23(1):103{145, 2005.

[5] S. S. Anand and B. Mobasher. Contextual recommendation. In Lecture Notes In Arti_cial Intelligence, volume 4737, pages 142-160, Springer-Verlag, Berlin, Heidelberg, 2007.

[6] Adomavicius, G., Tuzhilin, A.: Context-Aware Recommender Systems, chap. 7, pp. 217–253. Springer US, 2011.

[7] Campos, P., Fernández-Tobías, I., Cantador, I., Díez, F. "Context-Aware Movie Recommendations: An Empirical Comparison of Pre-Filtering, Post-Filtering and Contextual Modeling Approaches". In EC-WEB, 2013.

[8] L. Baltrunas and F. Ricci. Context-based splitting of item ratings in collaborative filtering. In Proceedings of ACM conference on Recommender systems, pages 245–248, 2009.

[9] Y. Zheng, R. Burke and B. Mobasher. "Splitting Approaches for Context-Aware Recommendation: An Empirical Study". Proceedings of the 29th ACM Symposium on Applied Computing, Gyeongju, South Korea, pp. 274-279, 2014.

[10] L. Baltrunas and F. Ricci. Experimental evaluation of context-dependent collaborative filtering using item splitting. User Modeling and User-Adapted Interaction, pages 1–28, 2013.

[11] Said, A., De Luca, E.W., Albayrak, S.: Inferring contextual user profiles-improving recommender performance. In: Proceedings of the 3rd RecSys Workshop on Context-Aware Recommender Systems, 2011.

[12] Sinha R, et al , Sinha R, Swearingen K., Comparing recommendations made by online systems and friends. In: Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries, 2001.

[13] Jamali and Ester, TrustWalker: A Random Walk Model for Combining Trust-based and Item-based Recommendation, SIGKDD, 2009.

[14] Jamali and Ester, A Matrix Factorization Technique with Trust Propagation for Recommendation in Social Networks, RecSys, 2010.

[15] Guo, G., Zhang, J., & Thalmann, D., Merging trust in collaborative filtering to alleviate data sparsity and cold start. Knowledge-Based Systems, 57, 57-68, 2014

[16] Azadjalal, M. M., Moradi, P., Abdollahpouri, A., & Jalili, M., A trust-aware recommendation method based on Pareto dominance and confidence concepts. Knowledge-Based Systems, 116, 130-143, 2017.

[17] Moradi, P., & Ahmadian, S., A reliability-based recommendation method to improve trust-aware recommender systems. Expert Systems with Applications, 42(21), 7386-7398, 2015.

[18] Parvin, H., Moradi, P., & Esmaeili, S., TCFACO: Trust-aware collaborative filtering method based on ant colony optimization. Expert Systems with Applications, 118, 152-168, 2019.

[19] Yadav, S., Kumar, V., Sinha, S., & Nagpal, S., Trust aware recommender system using swarm intelligence. Journal of computational science, 28, 180-192, 2018.

[20] Wang, M., Wu, Z., Sun, X., Feng, G., & Zhang, B., Trust-Aware Collaborative Filtering with a Denoising Autoencoder. Neural Processing Letters, 1-15, 2018.

[21] Li, Jiyun, Caiqi Sun, and Juntao Lv. "TCMF: trust-based context-aware matrix factorization for collaborative filtering." Tools with Artificial Intelligence (ICTAI), IEEE 26th International Conference on. IEEE, 2014.

[22] Baltrunas, Linas, Bernd Ludwig, and Francesco Ricci. "Matrix factorization techniques for context aware recommendation." Proceedings of the fifth ACM conference on Recommender systems. ACM, 2011.

[23] Li, Jiyun, Rongyuan Yang, and Linlin Jiang. "DTCMF: Dynamic trust-based context-aware matrix factorization for collaborative filtering." Information Technology, Networking, Electronic and Automation Control Conference, IEEE. IEEE, 2016.

[24] Keikha, Fateme, Mohammad Fathian, and M. Gholamian. "TB-CA: A hybrid method based on trust and context-aware for recommender system in social networks." Management Science Letters 5(5), 471-480, 2015.

[25] Xu, Jun, et al. "Trust-based context-aware mobile social network service recommendation." Wuhan University Journal of Natural Sciences 22.2: 149-156, 2017.

[26] Otebolaku, Abayomi, and Gyu Myoung Lee. "A Framework for Exploiting Internet of Things for Context-Aware Trust-Based Personalized Services." Mobile Information Systems 2018.

[27] Pedro G. Campos, Iván Cantador, Fernando Díez, Ignacio Fernández-Tobías: A Criterion Based on Fisher's Exact Test for Item Splitting in Context-Aware Recommender Systems. SCCC: 80-82, 2014.

[28] Yang, Shuxin, Qiuying Peng, and Le Chen. "The BPSO Based Complex Splitting of Context-Aware Recommendation." International Symposium on Intelligence Computation and Applications. Springer, Singapore, 2015.

[29] Zheng, Yong, Bamshad Mobasher, and Robin Burke. "Emotions in context-aware recommender systems." Emotions and Personality in Personalized Services. Springer, Cham, 311-326, 2016.

[30] Fan, Zi-Wei, et al. "Empirical evaluation of contextual combinations in recommendation system." 2016 International Conference on Machine Learning and Cybernetics (ICMLC). Vol. 2. IEEE, 2016.

[31] Fridi, Asmaa, and Sidi Mohamed Benslimane. "Towards Semantics-Aware Recommender System: A LOD-Based Approach." International Journal of Modern Education and Computer Science, 9(2), 55-61, 2017

[32] Herlocker, Jonathan L., et al. "An algorithmic framework for performing collaborative filtering." 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 1999.

[33] Massa, Paolo, and Paolo Avesani. "Trust-aware collaborative filtering for recommender systems." OTM Confederated International Conferences" On the Move to Meaningful Internet Systems". Springer, Berlin, Heidelberg, 2004

Mutual Progress of Big Data and Interactive Visualization

1st Moustafa Sadek KAHIL
Lamis Laboratory, Mathematics and Computer Science Department
Tebessa University
Tebessa, Algeria
moustafa-sadek.kahil@univ-tebessa.dz

2nd Abdelkrim BOURAMOUL
MISC Laboratory,
Fundamental Computer Science and its Applications Department,
Constantine2 University
Constantine, Algeria
abdelkrim.bouramoul@univ-constantine2.dz

3rd Makhlouf DERDOUR
Lamis Laboratory, Mathematics and Computer Science Department
Tebessa University
Tebessa, Algeria
m.derdour@yahoo.fr

Abstract— **The Big Data technology marks a wide growth in the different domains. With its major features, namely volume, variety and velocity, it has revealed many challenges, one of which is the data analysis and exploration. The data visualization is widely used for this purpose. However, with the Big Data characteristics, using the conventional visualization techniques and tools is nugatory. That's why new data visualization approaches and techniques must be thought. In this paper, we take an overview on the Big Data trending platforms and tools and their usage in the different challenges of this technology, especially the data storage and processing. We also present the data visualization techniques, as well as the libraries, services and platforms that are used to perform the data graphical presentation, while listing their limitations in the Big Data field. We mention the different solutions that are proposed in this regard. We finish by proposing an idea that aims to develop an approach for the data hierarchical visualization, a widespread technique in the Big Data field.**

Keywords— Big Data, NoSQL, MapReduce, HDFS, Apache Spark, Interactive Visualization techniques

I. INTRODUCTION

Big Data is the result of the almost complete dependence of computer by the individuals as well as the enterprises. This caused a rapid growth of data that come from different sources and reveal, therefore, many issues in regard with the data storage, processing, extraction and so on[1]. On the other hand, the data visualization is increasingly needed in order to simplify and enhance to user the data analysis, exploration and research. However, with the apparition of Big Data, other challenges regarding the data visualization were introduced: the conventional visualization systems and approaches cannot support the data with such characteristics, their large size and variety and their need to be visualized in acceptable times. We present in the first section the most used platforms for the data storage and data processing namely NoSQL databases, Hadoop MapReduce and Apache Spark, and highlight their different components and their usage, followed by a comparative study between these two latter. In the second section, we present an overview on the data visualization techniques and tools, present thereafter the challenges of Big Data visualization, and list the proposed approaches that present solutions in this field. Finally, we present a global idea about proposing an approach that will encompass that data hierarchical visualization in the Big Data field.

II. BIG DATA PLATFORMS AND FRAMEWORKS

The Big Data challenges are not just about data storage problem, but they have to overcome other problems which concern processing issues such as extraction of information, decision making, ... To ensure the scalability of processing, two techniques can be used: Scale up and Scale out[2] as shown in Figure 1. Scale Out techniques consist on adding machine instances or processing nodes to the system. They use many more homogeneous clusters with identical nodes. While the of Scale Up techniques, they move from the use of small machines (in term of performance) to more efficient ones that are equipped with better processors and more storage memory.

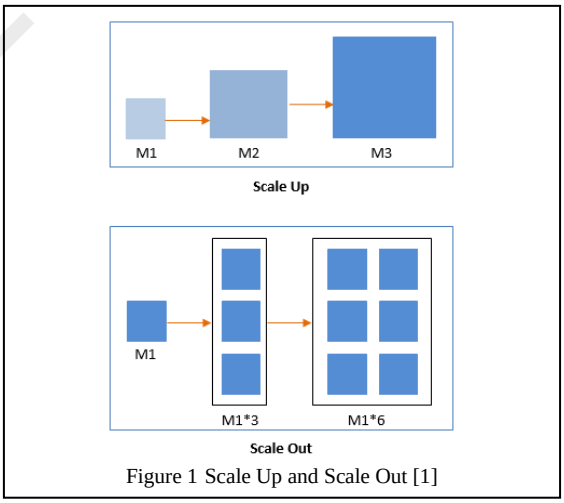


Figure 1 Scale Up and Scale Out [1]

A. Big Data Storage

With the data variety and huge size that mark the Big Data evolution, using the relational databases systems presents many limitations and becomes infeasible. The data storage is now performed by many mechanisms. We list the most relevant ones.

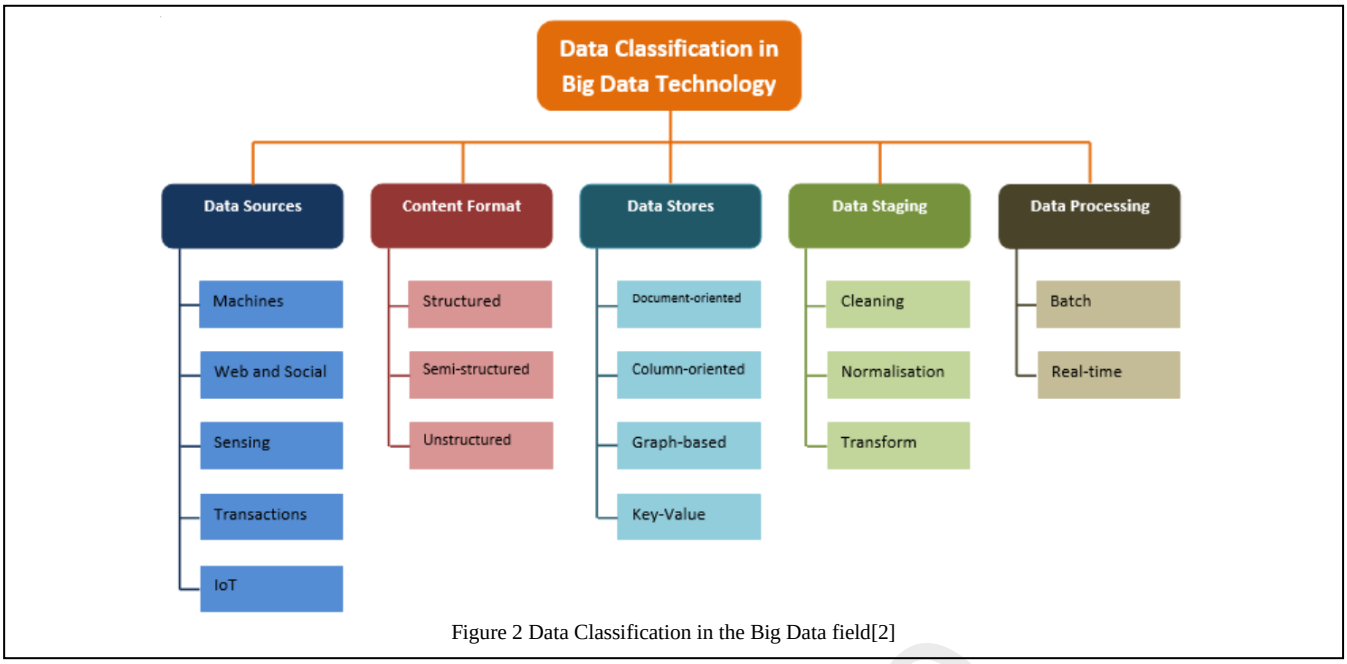


Figure 2 Data Classification in the Big Data field[2]

a) NoSQL Databases

NoSQL (Not only SQL) is the new Big Data storage databases. With the data variety, the storage must wrap up, in addition to structured data, other semi-structured and unstructured data that can be collected from different sources such as social networks, forums, sensors, etc. What reveals NoSQL databases.

These databases provide the mechanism to store and retrieve distributed data with large volume[3], [4]. They have several features compared to relational databases. We can cite[5]: (1) schema-free: unlike the relational databases, these databases may not be schematized a priori, i.e. the developer is not obliged to predefine all the tables' columns. Thus, any row of the same table does not necessarily have the same attributes as another, (2) ease of replication, (3) simple APIs, (4) Consistent and flexible operating modes, (5) Quick reading and writing operations, and (6) Large storage support.

NoSQL databases can be classified, as shown in Figure 2 [3], according to "Data Store" criteria, in:

- Column oriented databases: The principle of this type is to store and process the data according to columns other than lines (Example: Apache Cassandra). Hbase is a database management system that runs in Hadoop Distributed File System (HDFS). It belongs to this type of NoSQL databases it does not support SQL.
- Key-Value based databases: This category corresponds to simple data models based on unique key (Example: Redis, Couchbase Server).
- Oriented document databases: Their principle is similar to that of key-value based databases, except that documents can have different types including complex types (Example: MongoDB, CoucheDB).
- Graph based databases: They store and represent data using the graph model that includes nodes, arcs, and different graphs' relationships and properties (Example: Neo4J).

Other scientific communities (like the online educational and training site: www.edureka.co, for learning trending technologies) consider another class: Cache Systems based databases (Example: Redis, Memcache).

b) HDFS

HDFS is composed of a single name node and multiple data nodes [6] as shown in Figure 3[7]. The first type of node, which represents the master node, is responsible for: (1) access management to files by clients, (2) linking the name nodes with the data blocks, and (3) managing the file system operations such as opening, closing and renaming files and folders. Data nodes, composed of task tracker and job tracker, perform the various operations on data blocks such as creation, deletion and replication. The HDFS architecture involves storage in blocks of 64 megabytes. Each block is replicated in three copies: the second is stored in a local collection of data nodes called Local Rack, the third is stored in Remote Rack. The first type of node, which represents the master node, is responsible for: (1) access management to files by clients, (2) linking the name nodes with the data blocks, and (3) managing the file system operations such as opening, closing and renaming files and folders. Data nodes, composed of task tracker and job tracker, perform the various operations on data blocks such as creation, deletion and replication. The HDFS architecture involves storage in blocks of 64 megabytes. Each block is replicated in three copies: the second is stored in a local collection of data nodes called Local Rack, the third is stored in Remote Rack.

c) Data Center

The Data Center Physical Network is one of principal infrastructures of the Big Data storage. But this technology can aim farther: it can be a mechanism for acquiring, organizing and processing data, and even a source of data opportunity

(values and functions). It must also be admitted that the evolution and innovation in Data Centers (in capacity of processing and computing) are released due to the huge growth in Big Data applications. In [4], it was reported that Big Data has given Data Centers new features; it has made them take care, in addition to hardware facilitations, to strengthen the software side (acquisition, analysis, processing, organization...) in different areas and even to discover problems related to business operations and to find solutions for them.

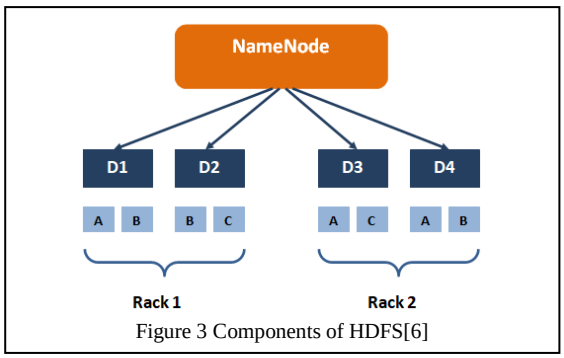


Figure 3 Components of HDFS[6]

In [8], many storage types that are used by many trending companies were shown. For example, Google and Facebook use respectively the NoSQL databases BigTable and Cassandra, Yahoo uses the cloud storage platform: Sherpa, and IBM uses Apache Hadoop and the platform Infosphere.

B. *Big Data Processing*

a) Hadoop

The famous Hadoop platform is the main reference to the Big Data technology. Apache Hadoop, developed at Yahoo by Doug Cutting in 2008, is a system designed to match with the large datasets that cannot be handled through conventional systems. Its principle is to distribute this massive data on machines so that these latter can process them. Its main advantage arises in its ability to support hundreds and even thousands of machines that can contain different configurations[7] and even different processor cores. Hadoop consists essentially of two 02 systems: (1) HDFS which is responsible of data storage (talked in section above), and (2) MapReduce, the charging of the parallel data processing through the distributed computing of clusters.

MapReduce is the major component of Hadoop which is responsible of the parallel processing of massive data. It is a simplified programming model that takes care of the parallelization task and allows programmers to focus on application development instead of thinking about doing this heavy task. It has two 02 essential functions for data processing which are: **Map** and **Reduce**. Its operation is indicated in Figure 4[6].

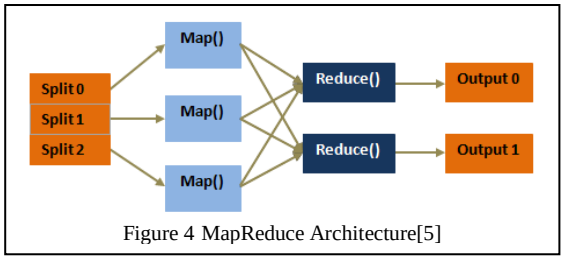


Figure 4 MapReduce Architecture[5]

In the Map phase, the data are outputted, after splitting them into fixed size blocks and passing them through the Mapper, as a key/value input pair. Here, they have become intermediate data[6]. In the Reduce phase, the Reducer groups the intermediate data with the same output key from the different functions of the Mapper. They are reduced into small sets of values, which produces the output.

b) Spark

Apache Spark is a system which is used to analyse massive data across groups of machines. It uses in memory computing which represents, according to [6], a high performance providing the most important functionality to speed up the processing of an application. Apache Spark has been developed to enhance Hadoop MapReduce functionalities[9]. In fact, it can process the data from Hadoop HDFS. It is simple to use, can run everywhere and combines: SQL, Complex analysis and Streaming. Retaining Hadoop MapReduce features such as fault tolerance and scalability, Apache Spark supports iterative algorithms for machine learning, Batch based applications and interactive data analysis tools[6].

Spark's basic data structure that serves these purposes is named: Resilient Distributed Datasets (RDD). It is an abstraction of distributed memory intended for programmers so that they can perform in memory calculation. RDDs are essentially depended on to serve for applications poorly served by other computing systems[10]: (1) Iterative algorithms and (2) Iterative data extraction tools. They provide a restricted mode for memory sharing to ensure fault tolerance.

c) Hadoop or Spark?

A question that must be asked before the process of developing a Big Data system: on which basis should we choose the appropriate platform and tools? For this, there are criteria through which we must make the comparison.

Following the comparative study done in [6], which was based on performances, we can resume the difference between Hadoop MapReduce and Apache Spark in these points: (1) MapReduce is more rapid in Batch processing, (2) it can directly process the data that are stored in the disk without loading them into the main memory, which is not the case in Spark (In-memory Computation), (3) it mainly deals with data sets that are formed beforehand, (4) Spark processes data only after loading them into the main memory, (5) it could be slow in the case of batch processing on large datasets, (6) it is mainly used for machine learning, broadcasting and batch processing.

Table 1 Comparison between Hadoop and Spark

Hadoop MapReduce	Apache Spark
• Batch processing • In-disc direct processing • Works with preformed datasets	• Real time processing and batch processing • In-memory processing • Used especially for machine learning and broadcast • Slow processing in batch mode

Thus, because MapReduce's processing is disk based, it can merge and partition mixed files and for fault tolerance it uses replication. While Spark, which is RAM based, it does not neither merge nor partition the mixed files and it uses RDDs for

fault tolerance. Table 1 summarizes the difference points, that we synthesized, between these two frameworks.

d) Important Hadoop and Spark Libraries

We cite here the most important libraries of Hadoop MapReduce and Apache Spark. Table 2 summarizes these libraries as well as their intended application.

Table 2 Important libraries of Hadoop and Spark

Framework	Libraries	Utility
Hadoop	Native Hadoop Library (Libhadoop.so)	Distributed processing of large data sets
	Libhdfs	Manipulate the HDFS files and the filesystem
Spark	MLlib[11]	For machine learning
	GraphX[12]	For graph processing
	Spark Streaming	For real-time processing
	Spark SQL	To structure data.
	Thunder	Analysing large-scale image and time series data

e) Other Frameworks

Yarn (Yet Another Resource Negotiator) is the next generation of MapReduce. This paradigm was developed to remedy the limitations of MapReduce, especially the FIFO scheduling where the real-time processing cannot be performed[7]. Likewise, many other frameworks were developed to enhance the real-time processing. For example, Apache Flink is dedicated to process stream data while supporting the event-time, processing-time and ingestion-time notions[13], Apache Kafka[14], based on the messaging mechanism, aims to solve the real-time problems in high throughput, and Apache Storm[15].

C. Data Formats

The data formatting allows to give to data a structure to so as simplify the different processing operations. Among the formats that are often used the Big Data field we find[16]:

- CSV (Comma Separated Values) formats data via separating the fields by commas, where every record is represented in on line.
- JSON (JavaScript Object Notation) is a superlative data exchange Format, because it is simple to be generated by machines and supported by the most of the current systems and programming languages. It includes two types of structural elements: (1) sets of "key"/"value" pairs and (2) lists of values. These elements comprise three data types:
- ORC (Optimized Row Compressed) is used to store the data in the Hive framework.

III. INTERACTIVE VISUALIZATION

Representing data to a machine (system) user can be defined as a resume of these data that is understandable to him. In particular, the statistical data; it is difficult or impossible to directly analyze data or extract information about a phenomenon, a study etc. without the help of the representation tools. The representation can be manifested in different forms: visual, acoustic, ... Visualization is a kind of data presentation of that consists to make data visual so that the target user understands them, i.e putting the results he seeks in graphical form. This is part of the human-computer interaction field. At first, the visualization consisted in representing users with static graphics that responded to their queries without intending to change or modify them according to their specific interests. For that, visualization has recently improved to interact with users by giving them privileges that allow them to customize the graphical representation according to their needs.

A. Data Visualization Techniques

There are several basic and advanced data visualization techniques. The most known basic techniques, that are presented in[17], [18] can be classified in four classes: Maps, Charts, Plots and Diagrams.

1) Maps techniques aim to visualize data by posinioning shapes on layout to specify a zone, object or element. Example of map presentations: Symbol Map.

2) Charts are generally used in the statistics to visualize the datasets by using lines, bar charts... The chart visualization can also be used to present data in real-time by using, for example, Streamgraph.

3) Plots are used to visualize many datasets together so as to view the relashionships between them, via scatters like Scatter Plots or bubbles like Circle Packing.

4) Diagrams can visualize various data types by hierarchical presentation such as Treemap and Circular Network Diagram, or multidimentional presentation such as Parallel Coordinate and Alluvial Diagrams.

Figure 5, Figure 6, Figure 7 and Figure 8 present respectively examples of each technique.

Likewise, there are many advanced visualization techniques such as the Virtual Reality and Augmented Reality.

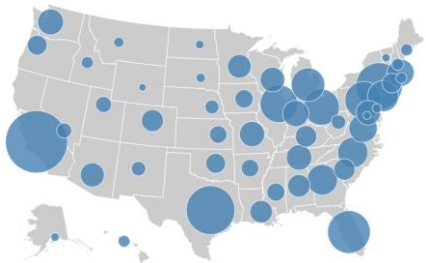


Figure 5 Symbol Map Example

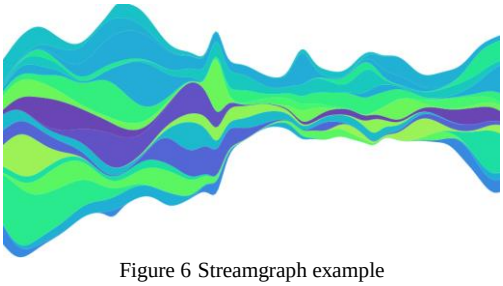


Figure 6 Streamgraph example

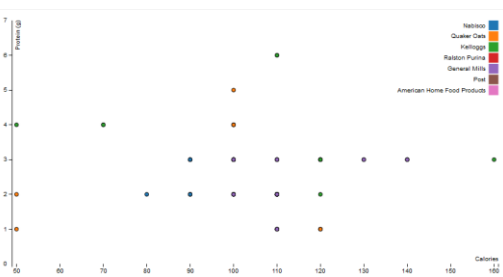


Figure 7 Scatter Plot example

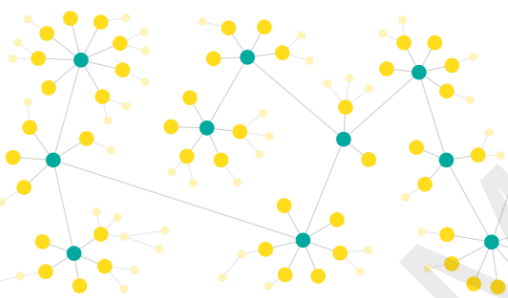


Figure 8 Circular Network Diagram example

B. Tools

Table 3 Examples of data visualization tools

Libraries	Vegas-Viz, D3js, React-vis, VX, Rechart, Flexmonster, Processingjs, Google Charts, Visjs, ZingChart, Gephi, Vega, RHadoop, Plotly
Platforms	Quire, Tableau Desktop, R, Infogram, IVEE, Weave, Microsoft Power BI, QlikView, Sisense
Services	CratoDB, Oracle Data Visualization, Tableau Server, Dundas BI, Looker

There are hundreds of visualization tools that are used in the data visualization. To choose the adequate tool, there is interest to compare them according to the criteria[18]: open source or not, capacity of integration with the popular data sources, interactivity, types of the clients (mobile, desktop, ...). We distinguish three types of data visualization tools[19]: libraries, platforms and services. Depending on these types, Table 3 illustrates each one.

C. Big Data Visualization

Data visualization has become a complex task in the Big Data filed. That is essentially because of the large size of data as well as their variety, the conventional visualization tools cannot, therefore, support these challenges. It has become necessary to develop approaches that aim to solve the Big Data visualizations problems in low latency while respecting a set of constraints namely: The interactive constraints to ensure the interactivity with users through different features such as zoom, pan, detail on demand[18]… The scalability constraints[19] to ensure the real-time scalability, the interactive scalability and the perceptual scalability. The structuring to optimize the different visualization tasks such as the data integration and the ease of operations[20]. To implement a visualization solution in a Big Data field, there is interest to refer to the classification that was mentioned in [21] to target the specific needs: visualize (1) a data type, (2) a dataset, or (3) a special topic. The first category concerns a particular subject during the visualization algorithm (visualization of the traffic of a network, visualization of the mails, ...). The second one concerns particular types of data (textual data visualization, video data visualization, etc.). And the third one concerns the visualization of particular datasets (a medical dataset visualization, a metrological dataset visualization, etc.).

The low latency is the most important challenge in the Big Data visualization. In this regard, several approaches were proposed. For example, in [22] and [23] consist on processing the complex query on cloud. In[24], the data are visualized progressively until the end of query processing. The proposed approach in [25] (VisReduce) aims to provide an interactive and scalable visualization by distributing the processing according to the MapReduce principle.

Other works aim to implement the data visualization techniques in a Big Data context. The Binned Aggregation technique was implemented in [22] for the healthcare domain. The parallel Coordinate technique was implemented in [21]. Circle Packing [26].

IV. Perspective on a promizing architecture for Big Data Hierarchical Interactive Visualization

The data hierarchical visualization is widespread in many Big Data important domains such as the medical and the educational domains. In this regard, we aim to propose an approach that will allow to ensure the data hierarchical visualization in the Big Data domains, while providing features to allow the interactivity with user so as he can personalize his own visualization and explore the data easily according to his needs, and respecting the visualization constraints. The approach must cover the hole visualization process from the data parsing to the graphical presentation. For that, an intuitive idea may be to propose a modular architecture that takes into consideration to provide a multilevel interactive visualization, where there are: a data parsing module to parse data from different sources and format them in a known format such as CSV and JSON, a criteria

handling module to organize the hierarchy levels' sequence, a visualization module to visualize the formatted data according to the existing criteria, and a data updating module to ensure the different update operations such as adding data sources and deleting the obsolete data. Figure 9 shows an overview of the potential architecture.

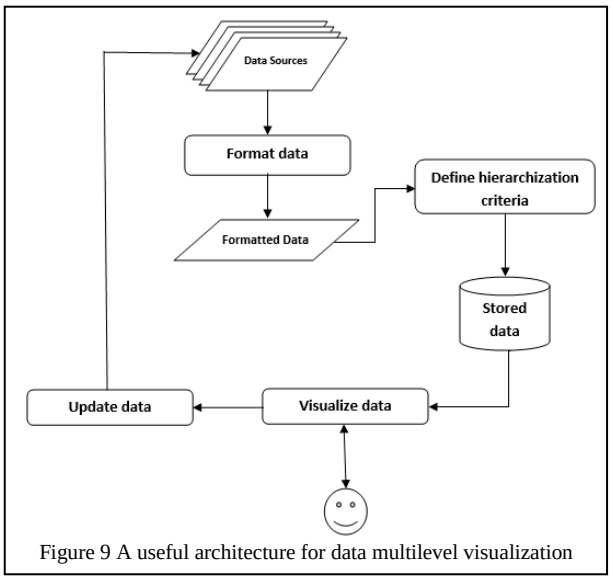


Figure 9 A useful architecture for data multilevel visualization

V. CONCLUSION

We presented by this paper the different concepts of Big Data as well as the data visualization, while highlighting their coevolution in terms of challenges, techniques and tools. The data visualization represents an important and unavoidable way to explore and analyze data in the different Big Data domains, so as to simplify the research process in the heterogenous and large sized data. We also highlighted the importance of the data hierarchical visualization in the Big Data field. In this regard, presented an idea to develop an approach that will aim to include the hole data hierarchical visualization process from the data parsing to the graphical presentation.

REFERENCES

[1] D. Rajasekar, C. Dhanamani, and S. K. Sandhya, 'A Survey on Big Data Concepts and Tools', vol. 5, no. 2, p. 5, 2015.

[2] K. Hwang, Y. Shi, and X. Bai, 'Scale-Out vs. Scale-Up Techniques for Cloud Performance and Productivity', in *2014 IEEE 6th International Conference on Cloud Computing Technology and Science*, Singapore, Singapore, 2014, pp. 763–768.

[3] I. A. T. Hashem, I. Yaqoob, N. B. Anuar, S. Mokhtar, A. Gani, and S. Ullah Khan, 'The rise of "big data" on cloud computing: Review and open research issues', *Inf. Syst.*, vol. 47, pp. 98–115, Jan. 2015.

[4] M. Chen, S. Mao, and Y. Liu, 'Big Data: A Survey', *Mob. Netw. Appl.*, vol. 19, no. 2, pp. 171–209, Apr. 2014.

[5] Jing Han, Haihong E, Guan Le, and Jian Du, 'Survey on NoSQL database', in *2011 6th International Conference on Pervasive Computing and Applications*, Port Elizabeth, South Africa, 2011, pp. 363–366.

[6] A. V. Hazarika, G. J. S. R. Ram, and E. Jain, 'Performance comparision of Hadoop and spark engine', in *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, Palladam, Tamilnadu, India, 2017, pp. 671–674.

[7] A. P. Kulkarni and M. Khandewal, 'Survey on Hadoop and Introduction to YARN', vol. 4, no. 5, p. 6, 2014.

[8] H. Fang, Z. Zhang, C. J. Wang, M. Daneshmand, C. Wang, and H. Wang, 'A survey of big data research', *IEEE Netw.*, vol. 29, no. 5, pp. 6–9, Sep. 2015.

[9] R. C. Maheshwar and D. Haritha, 'Survey on high performance analytics of bigdata with apache spark', in *2016 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)*, Ramanathapuram, India, 2016, pp. 721–725.

[10] M. Zaharia *et al.*, 'Resilient Distributed Datasets: A Fault-Tolerant Abstraction for In-Memory Cluster Computing', p. 14.

[11] X. Meng *et al.*, 'MLlib: Machine Learning in Apache Spark', p. 7.

[12] J. E. Gonzalez, R. S. Xin, A. Dave, D. Crankshaw, M. J. Franklin, and I. Stoica, 'GraphX: Graph Processing in a Distributed Dataflow Framework', p. 16.

[13] P. Carbone, A. Katsifodimos, S. Ewen, V. Markl, S. Haridi, and K. Tzoumas, 'Apache Flink™: Stream and Batch Processing in a Single Engine', p. 12.

[14] K. M. M. Thein, 'Apache Kafka: Next Generation Distributed Messaging System', p. 6.

[15] SZABIST, M. Hussain Iqbal, and T. Rahim Soomro, 'Big Data Analysis: Apache Storm Perspective', *Int. J. Comput. Trends Technol.*, vol. 19, no. 1, pp. 9–14, Jan. 2015.

[16] S. Ahmed, M. Usman, J. Ferzund, M. Atif, A. Rehman, and A. Mehmood, 'Modern Data Formats for Big Bioinformatics Data Analytics', *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 4, 2017.

[17] S. M. Ali, N. Gupta, G. K. Nayak, and R. K. Lenka, 'Big data visualization: Tools and challenges', in *2016 2nd International Conference on Contemporary Computing and Informatics (IC3I)*, Greater Noida, India, 2016, pp. 656–660.

[18] L. Wang, G. Wang, and C. A. Alexander, 'Big Data and Visualization: Methods, Challenges and Technology Progress', p. 6.

[19] R. Agrawal, A. Kadadi, X. Dai, and F. Andres, 'Challenges and opportunities with big data visualization', in *Proceedings of the 7th International Conference on Management of computational and collective intElligence in Digital EcoSystems - MEDES '15*, Caraguatatuba, Brazil, 2015, pp. 169–173.

[20] Y. Xu, W. Zhou, B. Cui, and L. Lu, 'Research on performance optimization and visualization tool of Hadoop', in *2015 10th International Conference on Computer Science & Education (ICCSE)*, Cambridge, United Kingdom, 2015, pp. 149–153.

[21] J. Zhang, M. L. Huang, W. B. Wang, L. F. Lu, and Z.-P. Meng, 'Big Data Density Analytics Using Parallel Coordinate Visualization', in *2014 IEEE 17th International Conference on Computational Science and Engineering*, Chengdu, China, 2014, pp. 1115–1120.

[22] Qunchao Fu, Wanheng Liu, Tengfei Xue, Heng Gu, Siyue Zhang, and Cong Wang, 'A big data processing methods for visualization', in *2014 IEEE 3rd International Conference on Cloud Computing and Intelligence Systems*, Shenzhen, China, 2014, pp. 571–575.

[23] B. Chandramouli, J. Goldstein, and A. Quamar, 'Scalable progressive analytics on big data in the cloud', *Proc. VLDB Endow.*, vol. 6, no. 14, pp. 1726–1737, Sep. 2013.

[24] C. D. Stolper, A. Perer, and D. Gotz, 'Progressive Visual Analytics: User-Driven Visual Exploration of In-Progress Analytics', *IEEE Trans. Vis. Comput. Graph.*, vol. 20, no. 12, pp. 1653–1662, Dec. 2014.

[25] J.-F. Im, F. G. Villegas, and M. J. McGuffin, 'VisReduce: Fast and responsive incremental information visualization of large datasets', in *2013 IEEE International Conference on Big Data*, Silicon Valley, CA, 2013, pp. 25–32.

[26] W. Wang, H. Wang, G. Dai, and H. Wang, 'Visualization of large hierarchical data by circle packing', in *Proceedings of the SIGCHI conference on Human Factors in computing systems - CHI '06*, Montréal, Québec, Canada, 2006, p. 517.

Tasks Scheduling Optimization for Scientific Workflow Application in Cloud Computing

1st Kouidri Siham
dept. of computer science
Faculty of exact and applied science,
University Oran1
Oran, Algeria
skouidri2008@gmail.com

2nd Yagoubi Belabbas
dept. of computer science
Faculty of exact and applied science,
University Oran1
Oran, Algeria
dyagoubi@gmail.com

***Abstract*—Cloud computing is a new era of network based computing and it allows users to pay as you need and has the high performance. Nowadays, task scheduling problem is the current research topic in cloud computing. In addition, a complex application can provide a very large amount of data and have a large number of tasks that may cause an increase in total cost of execution of that application, if not scheduled in an optimized way. In this paper in order to minimize the cost of the processing. We proposed a 2- stage tasks scheduling strategy. In the initial stage, we cluster the datasets based on their dependencies for a first scheduling, and optimize the tasks scheduling based on load virtual machine. Compared with previous work, simulation results show that our proposed strategy can effectively reduce the cost of the processing of the scientific workflow.**

Keywords—Cloud Computing, Scientific Workflow, Task Scheduling Optimization, Data Dependency.

I. INTRODUCTION

Cloud computing has emerged as a new computing platform to provide virtualized IT resources as cloud services by using the Internet technology [1]. It is a model, which provides easy access to available resources to cloud users on their demand [2]. Cloud computing enhances its performance and throughput by using an efficient task scheduling algorithm. Many scientific workflow applications in areas such as bioinformatics and astronomy require workflow processing in which tasks are executed based on their data dependencies. A scientific workflow is a specification of process to streamline automates and represents the schedule of integration, dataset selection, analysis and computation for the final presentation and visualization [3].

Scientific workflow applications are mainly used by scientists for their research purposes which are made up of coarse-grained and precedence constrained tasks. Scheduling scientific workflow tasks is the problem of scheduling tasks in scientific applications and mapping each task to suitable resources based on some performance impact factors [4].

The objective of this paper is to propose an optimization of scientific workflow tasks scheduling that minimize completion time based on data dependency and load of virtual machine. Our tasks scheduling approach takes place in two stages. Firstly, we schedule tasks based on data dependency and estimate completion time of task, on each virtual machine containing subset of datasets needed for this task, with actual load of virtual machine, after sorting the VMs based on the existence of datasets.

This paper is organized as follows: A few similar works have been discussed in Section II. The proposed scheduling approach has been explained under Section III. The evaluated results are given in section IV followed by conclusion in section V.

II. RELATED WORK

Scheduling of tasks is considered a critical issue in the cloud computing environment, A comparative study of task scheduling algorithms on the cloud computing environment has been done [5]:

• Round Robin: It is the simplest algorithm that uses the concept of time quantum or slices. Here the time is divided into multiple slices, each node is given a particular time quantum or time interval, and in this quantum, the node will perform its operations.

• Preemptive Priority: Priority of jobs is an important issue in scheduling because some jobs should be serviced earlier than other jobs that cannot stay for a long time in a system. A suitable job scheduling algorithm must be considered priority of jobs.

• The Shortest Job First (SJF): An SJF algorithm is simply a priority algorithm where the priority is the inverse of the next CPU burst. That is, the longer the CPU burst, the lower the priority and vice versa.

The emergence of cloud computing [6] offers a promising alternative for executing scientific workflows. In academia, the works [7-9] discussed the possibility and trade-offs of running scientific workflows on cloud.

In [10], a matrix based k-means clustering strategy for data placement has been proposed. It contains two algorithms that group the existing datasets into k data centers during the workflow build-time stageand allocate the newly generated datasets to the most appropriate data center, and has the largest dependency with them. This strategy can

effectively reduce the data movement of their workflow cloud systems. Their method aims to reduce the frequency of movement of the data, which does not mean reducing the load of the machines due to the execution of the tasks using the same datasets, which will cause a degradation of that performance.

III. Our Proposed Approach

Representation and scheduling of all the tasks in a data intensive application workflow is always a complex and critical job. A simple example of typical workflow is shown as Figure 1. Some tasks need more than one input dataset, while some datasets are required by more than one task. Placing these data items according to their dependencies can effectively reduce data transfers in cloud.

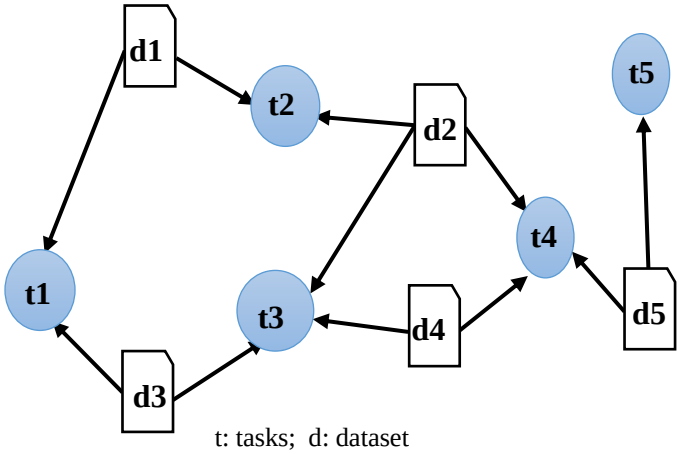


Figure 1. A simple instance of scientific workflow

Our proposed tasks scheduling for scientific workflow consists of two parts: first, we cluster the datasets of tasks based on their dependencies in the objective to schedule tasks to the virtual machine containing the majority of data required for each task. The second stage, we optimize the tasks scheduling based on load of virtual machine. These two stage are summarized as follows:

A. *Stage1: Tasks scheduling based on data clustering*

During this stage, we use a matrix model to represent the existing data [10]. We calculate the data dependencies of all the datasets and build up a dependency matrix DM. It can be calculated by counting the tasks in common between the task sets of d_i and d_j, which are denoted as T_i and T_j. Specially, for the elements in the diagonal of DM, each value means the number of tasks that will use this dataset (see figure 2).

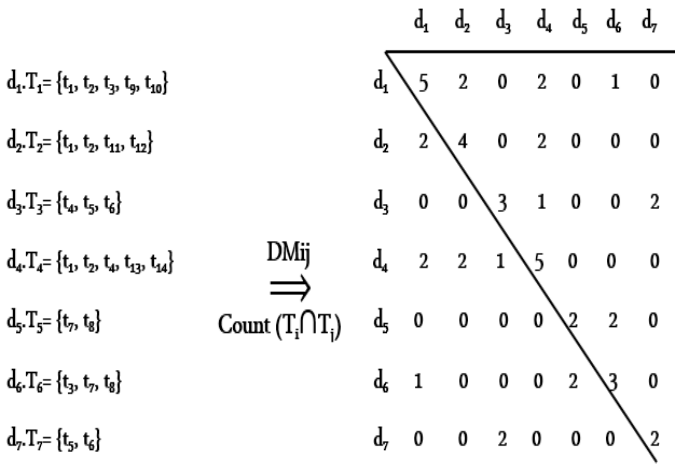


Figure 2. Dependency Matrix

Next, we use the BEA (Bond Energy Algorithm) [11] to transform the dependency matrix DM to clustering matrix CM. In CM, the items with similar values are grouped together (i.e. large values with other large values, and small values with other small values). Fig 3; show the CM of the example of DM after BEA transformation.

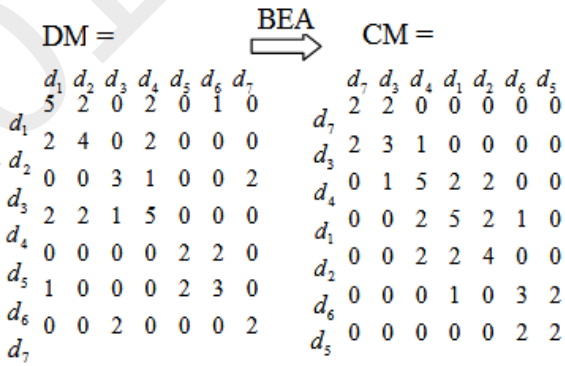


Figure 3. Transformation DM to CM by BEA

After the above BEA operation, recursive partition will be performed. For each partition of the matrix, a division position p is determined to maximize the following objective function in formula 1 [12].

$$CR = \left(\sum_{i=1}^{p}\sum_{j=1}^{p} C_{ij} + \sum_{i=p+1}^{n}\sum_{j=p+1}^{n} C_{ij}\right) / \left(\sum_{i=1}^{p}\sum_{j=p+1}^{n} C_{ij}\right) \quad (1)$$

This measurement, CR, means that the datasets in each partition have higher dependencies with each other and lower dependencies with the datasets in the other partitions. Based on this measure we can simply calculate all the CRs for p=1,2...n−1, and choose p such that it has the maximum CR value as the partition point.

In the situation of Fig. 3, the values of CR are calculated on CM as follows: 22, 46, 10, 12.66, 46, 22 respectively CR1, CR2, CR3, CR4, CR5, CR6. We can deduce that the best value of p = 2. After this partition, the CM forms two new clustered matrices. we denote the top one as CM_T, which contains the dependencies of datasets $D_T=\{d_1,d_2...d_p\}$and the bottom one as CM_B, which contains the dependencies of datasets

$D_B= \{d_{p+1},d_{p+2}...d_n\}$.

Algorithm1: Partitioning Algorithm of CM

1. **Input :** CM (Clustering Matrix)
2. **Output :** CM_T , CM_T

Description

3. **for** p=1 ; p<n-1 ; p++ **do**

   Calculate CR[p]; (formula 1)
4. **Choose** p / CR[p]=max
5. **Partitioning** CM into CMh et CMd // depending on the partition point p
6. **Return** CM_T, CM_T;

End

After numerous iterations, the result of the example will give the result as show in figure 4.

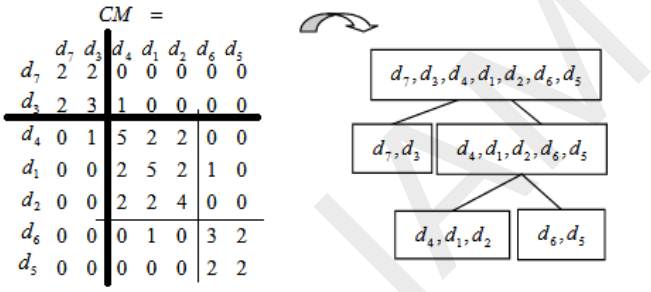


Figure 4. Partitioninf of CM

We will distribute clustered datasets to virtual machines (VMs), this part is designed to assign dataset groups to existing virtual machines where their initial storage capacity is calculated, sorted and verified if they can host the most suitable and an ideal storage datasets group.

Recursive calls to the distribution algorithm are executed until the end of the last partition.

Algorithm2: Distribution of datasets

Input: datasets, VmList

Output: allocation des datasets

Description

1. Sort VM according to the free storage space;
2. For each $dataset_i$
3. S_i=Size($dataset_i$)
4. Select the VM with enough storage space to allocate dataseti

End

B. Stage2: Optimization of tasks scheduling

- At first, the technique used is based on the placement of datasets, that is, the ready tasks are scheduled to the virtual machine (VM) that contains most of the required datasets [10]. The problem of this approach that did not take into account the load of the virtual machine (VM), if each task runs in the VM that contains the majority of the datasets, we can get the situation where the machine is running the majority of tasks with same of datasets that will cause a load of such a machine and a degradation of their performance.
- The goal of our approach is to reduce the response time of tasks by estimating the response time of each one of them with the current load of the virtual machine. The decision to reschedule the task is performed, if only, the processing time of the task with its local data exceeds the processing time with the same data transferred to another machine.

  The estimate of the response time of each task is calculated according to the following formulas:

$$RT_{ti} = T_{exec_i} + Data\_access\_ti + T\_release \quad (2)$$

$$Data\_access\_ti = \sum_{k=1}^{n} \frac{local_{Dataset_{Size_k}}}{Disc-transfert-capacity_{VMj_{ti}}} + Remote\_DA \quad (3)$$

$$T\_release = T\_release\_processor\_VMj + T\_release\_Dataset \quad (4)$$

Where:

RT_ti: Response Time of task i;

Data_access_ti: Access Time to data (local and remote)

T_release: resource release time (processors and dataset)

IV. Performance Evaluation

To validate the proposed approach we have implemented our algorithm in CloudSim[13], The CloudSim simulation layer provides support for modeling and simulation of virtualized Cloud-based data center environments including dedicated management interfaces for virtual machines (VMs), memory, storage, and bandwidth. This layer handles the fundamental issues such as provisioning of hosts to VMs, managing application execution, and monitoring dynamic system state.
So to study our approach using the CloudSim, we proposed a simulation environment that has the following parameters: 2 data centers are created, the service providers are represented by 100 virtual machines, and the processing elements (PEs) number of each virtual machine is within the range of 2 to 4. 200 tasks are submitted to the service providers.

We run each workflow instance through 3 simulation strategies:
Random: In this simulation, we place randomly the existing data in the Cloud.
Building Time: This strategy consists< to place the data after the clustering and distributed algorithm.
Optimazing Tasks Scheduling (Proposed Approach): This simulation shows the performance of our proposed algorithm.

A. Reponse Time vs Number of cloudlets

We have measured our approach (PA) with Build Time strategy [10] and random placement of datasets approach. We note that our approach minimizes the response time compared with Build Time and Random strategies (see fig 5).

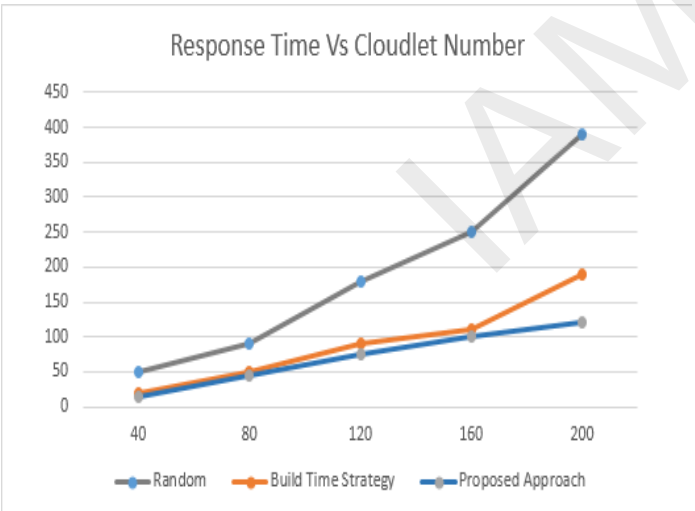


Figure 5. Response Time Vs Number of Cloudlet

B. Reponse Time Vs Number of VMs

In this simulation, we set the file size and vary the number of Vms (10, 20, 30, 50, 100) and see the impact on the execution time, we notice a decrease in the execution time compared to other strategies.

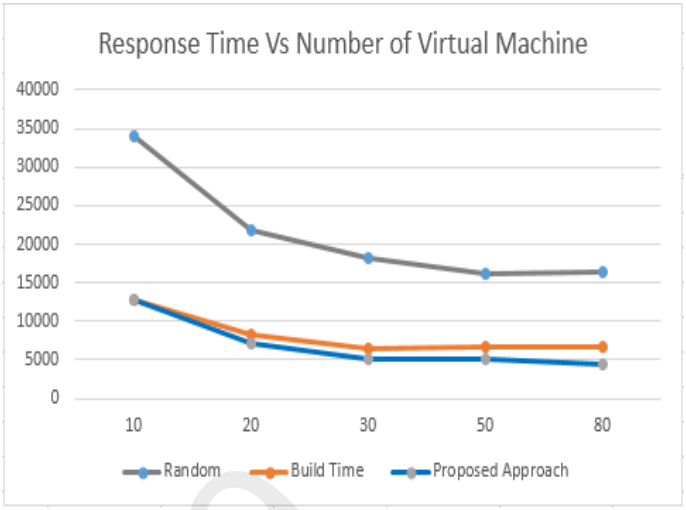


Figure 6. Response Time Vs Number of Virtual Machine

V. Conclusion

Cloud computing is emerging as one of the best solution to utilize existing resources for catering the massive computational and data handling requirements of today's high performance applications. The Cloud storage provides the ability to access and manage data and data resources on the Cloud.
In this paper, we have proposed an optimizing tasks scheduling for scientific workflow application in cloud computing, based on two stage.
The simulation result showed that our proposed strategy give the better response time compared with Build Time [10] and random strategies.
In future work, we will use a real platform to implement our proposition.

VI. Reference

[1] B. Hayes, "Cloud computing", Communications of the ACM, vol. 51, no. 7, 2008, pp. 9-11.

[2] D. a. S. N. R.Nallakumar, "A Survey of Task Scheduling Methods in Cloud Computing," International Journal of Computer Science and

[3] Engineering (JCSE), vol. 2, no. 10, 31 Oct 2014.

[4] SaeidAbrishami,MahmoudNaghibzadeh and Dick H.J Epema," Cost-Driven Scheduling of Grid Workflows Using Partial Critical Paths", IEEE Transactions on Parallel and Distributed Systems, Vol.23, No.8, Aug 2012.

[5] Vignesh,V., Kumar,K.S., Jaisankar,N.: "Resource management and scheduling in cloud environment", International Journal of Scienti_c and Research Publications, vol. 3, p. 1, (2013).

[6] A. Weiss, "Computing in the clouds". netWorker, 11(4):16–25, 2007.

[7] E. Deelman, G. Singh, M. Livny, B. Berriman, and J. Good, "The cost of doing science on the cloud: the montage example," Proceedings of the ACM/IEEE conference on Supercomputing, 2008.

[8] C. Hoffa, G. Mehta, T. Freeman, E. Deelman, K. Keahey, B. Berriman, and J. Good, "On the use of cloud computing for scientific workflows,"

3rd International Workshop on Scientific Workflows and Business Workflow Standards in e-Science in conjunction with Fourth IEEE International Conference on e-Science, 2008.

[9] Ü. V. Çatalyürek, K. Kaya, B. Uçar, "Integrated Data Placement and Task Assignment for Scientific Workflows" in Clouds.DIDC' 11, June 8, 2011, San Jose, California, USA.

[10] D. Yuan, Y. Yang, X. Liu, and J. Chen, "A data placement strategy in scientific cloud workflows," J. Future Generation Computer System, 26(8): 1200-1214, 2010.

[11] W.T. McCormick, P.J. Sehweitzer, T.W. White, Problem decomposition anddatareorganizationbyaclusteringtechnique,OperationsResearch20(1972)993–1009.

[12] Qing,Z., Congcong, X. ,Kunyu,Z., Yang,Y., Jucheng,Y.: PA Data PlacementAlgorithm for Data Intensive Applications in Cloud, International Journal of Gridand Distributed Computing, Vol. 9, No. 2, pp.145-156, 2016.

[13] Calheiros,R., Ranjan,R., Beloglazov1,A., De Rose,C., Buyya,R. :CloudSim: a toolkit for modeling and simulation of cloud computing environmentsand evaluation of resource provisioning algorithms, Software: Practice and Experience journal, 41(1): 23{50, 2011.

Bi-objective flowshop scheduling under unavailability constraints for minimizing total tardiness and peak power consumption

Radhwan Boufellouh
Manufacturing Engineering Laboratory
University of Tlemcen
Tlemcen, Algeria
ra_boufellouh@enst.dz

Fayçal Belkaid
Manufacturing Engineering Laboratory
University of Tlemcen
Tlemcen, Algeria
f_belkaid@yahoo.fr

***Abstract*— Non Dominated Sorting Genetic Algorithm (NSGA-II) is adopted to deal with a multi-objective flow shop scheduling problem. In particular, we consider a non-permutation flow shop scheduling problem under machine unavailability constraints with objectives of finding good compromise between total tardiness of jobs and peak power consumption. We propose a solution encoding scheme that facilitates the search through the vast solution space, then, assess the performance of the solving algorithm and find the best parameter values that yields the best results.**

Keywords*— Peak power load, Flowshop scheduling, Unavailability constraints, Total tardiness, Bi-objective optimization, NSGA-II

I. INTRODUCTION

The industrial sector currently accounts for most of the worldwide total energy consumption [1]. The consumption of energy by the sector has almost doubled over the last 60 years [2]. Industrial energy consumption is expected to increase 40% from 175 quadrillion Btu in 2006 to 246 quadrillion Btu by the end of the next decade [3]. Furthermore, the high economic growth in Africa leads to an expansion of the manufacturing sector and an increase in its industrial energy consumption by 30% in 2040 [1].

Motivated by the aforementioned global statistics, there has been recently a growing interest in green manufacturing with considering energy saving as one of the crucial objectives. This research attempts to tackle this problem from a system level perspective by exploring the use of scheduling as a way to temper power consumption. A moderate amount of research has been built that focuses on investigating scheduling problems with power down mechanism of machines [e.g. 5, 6, 7, 9, 10]

Restraining peak power load is an emergent approach in demand-side management due to the decisive role it plays in the energy costs of large electricity-demand manufacturing enterprises. [4] considered a permutation flow shop scheduling problem with variable processing speeds to minimize the makespan under a non-classical constraint on peak power load. [5] approached a flow shop scheduling problem with limited intermediate storage under peak power load upper bound by adapting a mixed-integer nonlinear programming as a solving method. [6] developed a mixed integer linear programming model for the flexible flow shop scheduling problem, subject to peak power load constraint, with the objective of minimizing the weighted sum of the total tardiness and the makespan.

It seems from the relevant literature that the general trend in approaching the problem is by imposing a hard constraint on the maximum power consumption, then, solving the problem by minimizing the classical scheduling objective, which happened to be in most cases the makespan of the schedule. In our attempt, we will take a different paradigm, by solving the problem using a bi-objective approach where the peak power consumption minimization is one of the objectives alongside with minimizing total tardiness which is a very important objective for manufacturing enterprises since it links production scheduling directly to the delivery deadlines (also known as due dates) negotiated with customers. And so the underlying problem is to find a reasonable customer/producer compromise balancing between satisfying customer-imposed deadlines and reducing peak power consumption costs for the producer.

Due to the nature of energy costs in manufacturing environments, we address the multi-objective problem in a non-permutation flow shop under unavailability constraints. The unavailability periods represents idle time intervals where machines are shut down for already-scheduled maintenance interventions. In particular, we search for Pareto optimal schedules, or schedules for which no other schedule has both lower tardiness costs and lower peak power consumption. Only the non-preemptive and non-resumable cases are considered. i.e., once a job is released for processing on a particular machine, it must be processed completely without interruption by another job or by an unavailability period.

Contrary to articles [4] and [7], In order to handle the bi-criteria nature of this problem, we use NSGA-II as a combinatorial approach to approximate the Pareto optimal front and minimize the peak power consumption total tardiness simultaneously. We consider the problem when only one level of processing speeds is available. In addition, we consider flow shops with unlimited intermediate storage between machines.

The remainder of this paper is structured as follows: At first, the scheduling problem is formally described and a mathematical programming formulation is provided in section 2. Section 3 then explains the solving method implemented to solve the problem. Samples of results obtained from computational experiments are analyzed in section 4. At last, concluding remarks and future research scopes are provided in section 5.

II. Problem Discription

The problem treated in this paper can be described as follows: we are given a set of *n* jobs to be processed on a set of *m* machines in a flow shop routing environment. Each job *j* (*j* = 1,…,*n*) has a fixed processing time P_{ij} a power requirement per unit of time φ_{ij} when treated by machine (*i* = 1,…,*m*) and a due date d_j. The machines are subject to predetermined unavailability periods and each job has to be processed non-preemptively on each machine and not allowed to be resumed if its processing is interrupted by an unavailability period of a particular machine. Jobs are allowed to change their processing sequence between machines; therefore, an unlimited buffer size between machines is assumed. The objective is to find a reasonable approximation of the real Pareto front between total tardiness and the peak power demand.

To develop a mixed integer linear programming model for the problem, the following decision variables are defined: X_{ijt} = 1 if job *j* is processed on machine *i* at time *t*, and 0 otherwise. Y_{ijt} = 1 if job *j* is not yet completely processed on machine *i* by time *t*, and 0 otherwise. *Qmax*: The schedule's peak power consumption. C_{mj} is the completion time of job *j* and T_j is the tardiness of job *j*.

With the decision variables defined above, the investigated problem can be formulated as the following MILP model:

$$\min. Qmax \quad (1)$$

$$\min. \sum_{j=1}^{n} T_j \quad (2)$$

s. t.

$$\sum_{t=1}^{Cmax} X_{ijt} = P_{ij} \quad \forall i = 1,\dots,m \quad \forall j = 1,\dots,n \quad (3)$$

$$\sum_{j=1}^{n} X_{ijt} \le a_{it} \quad \forall i = 1,\dots,m \quad \forall t = 1,\dots,Cmax \quad (4)$$

$$\sum_{l=t}^{Cmax} X_{ijl} \le \left(1 - X_{i+1,jt}\right)P_{ij} \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m-1 \quad \forall t = 1,\dots,Cmax \quad (5)$$

$$Y_{ijt} \le 1 - X_{ij,t-2} + X_{ij,t-1} \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m \quad \forall t = 3,\dots,Cmax \quad (6)$$

$$Y_{ijt} \le Y_{ij,t-1} \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m \quad \forall t = 2,\dots,Cmax \quad (7)$$

$$Y_{ijt} = 1 \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m \quad \forall t = 1,2 \quad (8)$$

$$X_{ijt} \le Y_{ijt} \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m \quad \forall t = 1,\dots,Cmax \quad (9)$$

$$\sum_{j=1}^{n}\sum_{i=1}^{m} X_{ijt}\varphi_{ij} \le Qmax \quad \forall t = 1,\dots,Cmax \quad (10)$$

$$C_{mj} \ge tX_{ijt} \quad \forall j = 1,\dots,n \quad \forall t = 1,\dots,Cmax \quad (11)$$

$$T_j \ge C_{mj} - d_j \quad \forall j = 1,\dots,n \quad (12)$$

$$Qmax \ge 0 \quad (13)$$

$$X_{ijt}, Y_{ijt} \in \{0,1\} \quad \forall j = 1,\dots,n \quad \forall i = 1,\dots,m \quad \forall t = 1,\dots,Cmax \quad (14)$$

The objectives of the model are to minimize the peak power *Qmax* and the total tardiness of the schedule across the allowed makespan *Cmax*. Constraints (3) require that the total sum of assigned processing times slots of a job across the entire makespan is exactly equal to its processing time. Constraints (4) ensure that two operations on the same machine are not allowed to overlap and that no job is processed while the machine is unavailable. The parameter a_{it} takes value 1 if machine *i* is available at time *t* and 0 otherwise. Constraints (5) guarantee that a job cannot be processed on a particular machine unless it is fully completed on the previous one. Constraints (6)-(9) ensures that every job is processed non-preemptively on every machine. Constraints (6) require that a job is not allowed to be processed at instant *t* if its processing is stopped at the previous one. Constraints (7) ensure that once a job's processing is halted at a particular time instant it cannot be resumed again at the next one and consequently through the rest of the time horizon. Constraints (8) allow for a job to be started or completed within the first two time slots; this set of constraints is required for the computations in constraints (7). Constraints (9) associate the previously described sets of constraints with the processing of jobs to obtain only feasible schedules. Finally Constraints (10) compute the peak power of the schedule while (11) and (12) compute the tardiness.

III. Non-dominated sorting genetic algorithm II

The problem presented in this paper is strongly NP-hard since the single machine scheduling version with total tardiness as objective and where the machine is all the time available is already strongly NP-hard. Even for a small sized instance with 5 jobs and 3 machines, there are $5!^3 = 1728000$ possible way in which to sequence the jobs on machines. The machine unavailability disruptions along with the nonresumability requirement of jobs increases to the complexity of the problem by forcing the solution to use up the available time between consecutive maintenance interventions more effectively. Moreover, an interesting conflict will arise when jobs that (preferably according to their due date and tardiness weight characteristics) ought to be scheduled concurrently during a certain period of time, are forced to be either delayed or scheduled more early in order to reduce peak power consumption. If one wishes to obtain more information about the possible alternative solutions yielding better compromise between the two criteria of total tardiness and peak power load, one has to resort to more sophisticated and efficient solving approaches. Using the ε–constraint approach by relying on the above mathematical programming model would be time consuming and unpractical for handling larger sized problem

like ones encountered in real world industry situations, which has been observed by several researchers who have developed time indexed MILP's similar to ours. Thus, to be able to find a decent approximation of the Pareto frontier that characterizes the tradeoff between the two objectives, we resort to a multi-objective approach. Precisely, we adapt and implement the well-known Non-dominated Sorting Genetic Algorithm (Second Version) which is one of the most prominent population based metaheuristics to solve multi-objective scheduling problems. In what follows, we will detail how a particular solution of the problem is encoded so that the algorithm is capable of generating the complete schedule. Then, we will explain the various genetic operators, elitism and diversity mechanism involved in the evolutionary process of the population towards, hopefully, a global optimum.

A. Algorithm description

NSGA-II is a multi-objective evolutionary algorithm developed initially by Deb et al. [9]. The algorithm uses the concepts of dominance, crowding and elitism to guide the evolution of the population towards the Pareto optimal frontier and produce multiple diverse solutions in one single run. In NSGA-II, the population of solutions, composed of parents from the previous generation and newly generated children in the current generation by means of the genetic operators of crossover and mutation, is divided into several fronts. The non-dominated solutions receive rank 1 and form the first front; these solutions are then removed from the population and the new non-dominated solutions form the second front and receive rank 2 and so on until the total number of sorted solution surpasses half of the population's size. Individuals in the last front are further sorted according to their crowding distance if including all of them to the next generation's population is not possible because of the limit on the population size. The crowding distance of a solution is defined as the circumference of the hypercube defined by its neighbors, and infinity if it is a boundary point, i.e., a missing neighbor on at least one dimension of the objective space. Solutions with high crowding distance are considered better solutions, as they introduce more diversity in the population. At each iteration of the algorithm, an offspring population of size N is generated. The replacement phase of the population works as follows: the old and offspring populations are combined and ranked according to the previous two criteria: non-dominance and crowding. The better half of the union forms the new population. The selection mechanism used in our implementation of the algorithm is based on a binary tournament between two randomly sampled solutions from the population, then comparing them according to their respective non-dominance ranks and select the best individual. If the two picked solutions belong to the same front, the crowding criterion is then used to select the individual with the largest crowding distance. A more comprehensive description of the algorithm is presented in the paper published by Deb et al. [9].

In (Algorithm 1) The Non-dominated-sorting algorithm takes a population P of solutions and decomposes it into a sorted list of fronts F={F1, F2 ⋯} where |F| ⩽ |P|. The parameter ns is used to count the total number of solutions included in the set of fronts F and stop the sorting procedure once the number of solutions included equals or exceeds half of the population's size after forming and adding a particular front F_i which will then represent the last front on which the crowding distance assignment will be applied. It should be noted that Algorithm 1 can only be used for a bi-criteria case like ours. In-depth description and analysis of the algorithm is provided in the famous work by Talbi [8].

Algorithm 1 Non-dominated-sorting

Inputs:
$\mathcal{P}$ population

1: $N = |\mathcal{P}|$
2: $\mathcal{P}.sort\langle f_1, f_2\rangle$ — Sort $\mathcal{P}$ according to objective f_1 then according to objective f_2 if two solutions have the same f_1 value
3: $ns = 0$ — Initialize total added solutions
4: $i = 1$ — Initialize front counter
5: **while** $ns \leq N/2$
6: $p = \mathcal{P}.pull[1]$ — Pull out highest rank solution
7: $\mathcal{F}_i.add\{p\}$
8: $ns = ns + 1$ — Increment ns by one
9: **for each** $q \in \mathcal{P}$
10: **if** $(p \prec q) \| (p = q)$ **then**
11: $p = \mathcal{P}.pull\{q\}$ — Pull out q from $\mathcal{P}$
12: $\mathcal{F}_i.add\{p\}$ — q belongs to the i^{th} front
13: $ns = ns + 1$ — Increment ns by one
14: $i = i + 1$ — Increment front counter by one

B. Solution encoding

Instead of encoding the solution as a multi-field chromosome where each field contains a permutation of *n* numbers representing the actual sequence of jobs on each machine, we take advantage of the observation that in classical non-permutation flow shop problems the sequence of jobs does not dramatically change from one machine to another and most cases a few number of jobs actually switch positions in the sequences and often these jobs are adjacent. On account of this fact, we will rather use only one permutation field representing the actual sequence of jobs initially on M1 then use *m* - 1 continuous-allele parts to represent the change of positions between two consecutive machines and one more gene to scale the intensity of this change.

C. Genetic operators

As for the crossover operator, we use a two point crossover on the permutation part of the chromosome; this is explained in Fig. 1. The intermediate crossover is used in the rest of the chromosome where each parent will contribute with a fraction of its allele, while the second parent contributes with the rest to form the complete gene of the child; this is summarized in Fig. 2.

Parent 1	1	5	6	8	4	2	7	3

Parent 2	8	7	1	2	3	4	6	5

Offspring 1	1	5	8	2	4	6	7	3

Offspring 2	8	7	1	4	2	3	6	5

Fig. 1. Two point permutation crossover.

Parent 1	0.79	0.55	0.22	1.3
random 1	0.58	0.52	0.7	0.54
Parent 2	-0.95	0.19	0.8	0.09
random 2	0.44	0.8	0.89	0.19
Offspring 1	0.06	0.38	0.39	0.74
Offspring 2	0.02	0.26	0.74	1.07

Fig. 2. Intermediate crossover crossover.

Concerning the mutation operator, a two point mutation is implemented where, at first, two random points on the chromosome are selected. Then, the genes on the extremities of the two points are copied directly from one parent with their positions preserved. Finally, the missing genes are copied with their order of appearance in the second parent's chromosome. This procedure is illustrated in Fig. 3.

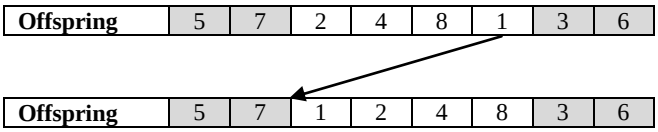


Fig. 3. Insertion mutation for permutation part.

Lastly, the mutation operator works on the continuous parts of the chromosome by changing their values randomly within their allowed ranges with the same probability for each gene. The range for the alleles of the permutation change is from -1 to 1. A negative value favors the possibility for the job in the corresponding position of the sequence to be moved earlier in the next machine's processing order, whereas a positive value increases the chance that the job will be pushed back later. The scaling gene allele's range is from 0 to n in order to reach all possible solution in the search space. A complete solution representation for a 3 machine 8 job example is presented in Fig. 4. The first row contains the job sequence on the first machine the second and third rows contains information about the position change of the jobs while the last gene in the third row represents the scaling gene. The procedure for obtaining the job permutations on machines $i = 2, \ldots, m$ is shown in Fig. 5.

1	5	6	8	4	2	7	3	
0.23	0.21	0.99	-0.98	-0.76	0.17	0.87	0.85	
-0.46	0	0.5	0.66	0.34	0.37	-0.2	-0.06	1.89

Fig. 4. Solution representation.

For obtaining the job sequences on all machines from the solution encoding presented above we first take the permutation presented in the first row as the job sequence on the first machine. Then, the sequence on machine M_i is obtained from perturbing the previous sequence on M_{i-1} for all $i = 2,\ldots,m$. If we consider our example in fig. 4, we should take the allele of each gene in the 2nd row and multiply it by the scaling gene's allele. Then, add the corresponding gene's position to obtain the values on the 2nd row in Fig. 5 below. For example the first value in the second row is equal to 0.23*1.89+1. After that, the job sequence on M1 is reorganized according to the rank of each gene (row 3 in Fig. 5) to obtain the job sequence on M2 and so on. For example, changing the position of job 6 from the 3rd position on machine M1 to the 2nd position on machine M2 is based on the sequence (M2r) obtained ranking the values in the previous row.

Pos.	1	2	3	4	5	6	7	8
	1.43	2.4	4.87	2.15	3.56	6.32	8.64	9.61
M2*r*	1	3	5	2	4	6	7	8
	0.13	2	3.95	5.25	5.64	6.7	6.62	7.89
M3*r*	1	2	3	4	5	7	6	8

1	5	6	8	4	2	7	3
1	6	4	5	8	2	7	3
1	6	4	5	8	7	2	3

Fig. 5. Job sequences' obtention for a three machine eight job example.

The algorithm is stopped once the population cease to evolve further beyond its current local optimum.

IV. Computational experimenttation

Computational experiments were conducted in order evaluate the performance of the metaheuristic under different probability values for crossover and mutation for different problem instances so that the parameters can be tuned. The algorithm was implemented on JAVA; all computations were performed on a 1.30-GHz Dual-Core PC with 2.00 Go of RAM. A set of 9 problems is proposed with each problem represented by the number of jobs n and the number of machines m. the set contains the following problems: (10, 2), (10, 5), (10, 10), (25, 2), (25, 5), (25, 10), (40, 2), (40, 5), (40, 10). To generate data parameters' for each problem instance, we used the following benchmark: the processing times and unavailability periods are generated from a uniform distribution U[1, 60]. The power demand of each job on each machine is from a uniform distribution U[1, 20]. Job due dates and start dates of machine unavailability periods are generated randomly in [0, Cmax] with Cmax = 30(n+m-1) and with, at most, three unavailability periods for each machine. The stopping criteria are 20 consecutive iterations without improvement in the Hyper-volume indicator, for the reason that it is a good predictor of convergence and diversity, or reaching a maximum of 15nm iterations. A population of 3nm chromosomes was chosen as an appropriate size for each run of the algorithm on each of the problems. It is known in population based metaheuristic literature that the initial population distribution has a tremendous effect on the performance and the progression of the search. To reduce the effect of the initial population

spread, save computational time required to perform the experiments and remove the need to resort to sophisticated diversification procedures, we use the same initial population for each run of the algorithm with different parameter values of crossover and mutation. The number of solutions obtained, the hyper-volume indicator of the obtained solution set and, the computational time taken to solve the problem were the three performance criteria chosen to measure the performance of the algorithm.

Fig. 6 and 7 presents the effect of the crossover and mutation probability values on the average number of obtained solutions for each algorithm parameter combination for each test problem. Fig. 8 presents a box plot of obtained number of solutions for each problem with the first problem is (10, 2) and the second is (10, 5) and so on according to the order of problems proposed previously. It can be noticed from the box plot depicted in Fig. 8 that the number of solutions is fairly stable with a small standard deviation of around 0.9 for the 10 jobs cases which is indicative of good convergence and robustness of the solving method. It is also obvious that the size of the non-dominated set is positively correlated with the size of the problem (both in terms of the number of jobs and the number of machines) as there more possibilities to arrange the jobs and obtain different solutions.

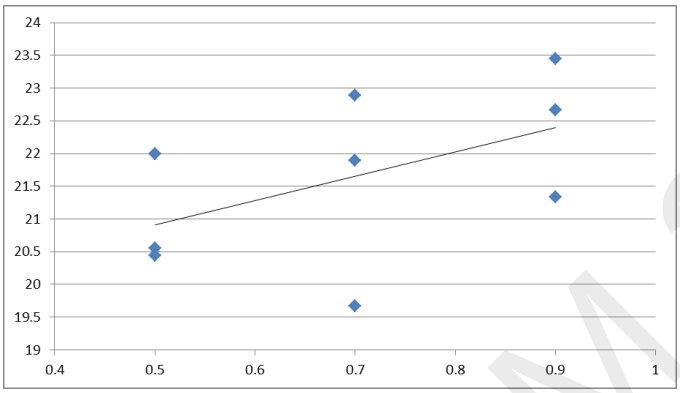


Fig. 6. Crossover probability value effect on average number of solutions.

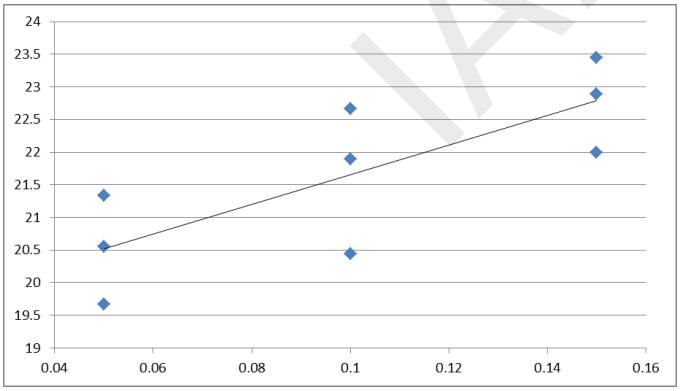


Fig. 7. Mutation probability value effect on average number of solutions.

Fig. 9 and 10 presents the effect of the crossover and mutation probability values on the values of the hyper-volume indicator of the solution sets obtained by each run of the algorithm for each test problem. Using the Wilcoxon signed-rank test to reveal the statistical significance of the difference between a two populations of hyper-volume values, one can infer that the best parameter value that yields maximum value for this measure is the 0.9 and 0.15 for the mutation probability value, i.e., the highest levels of these values. One can increase the mutation probability value beyond its limit to maybe obtain better solutions since its effect is far more visible than the effect of crossover.

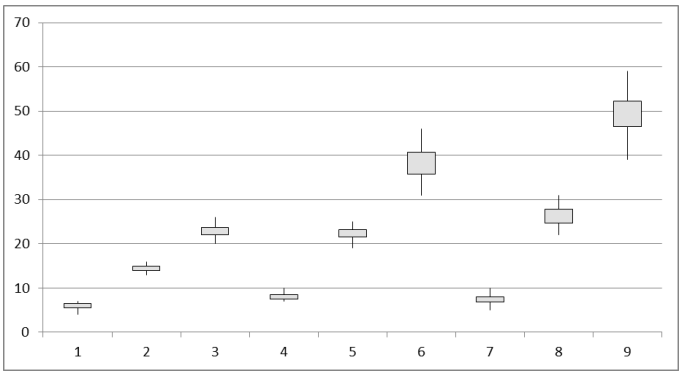


Fig. 8. Box plot of obtained number of solutions for each problem instance.

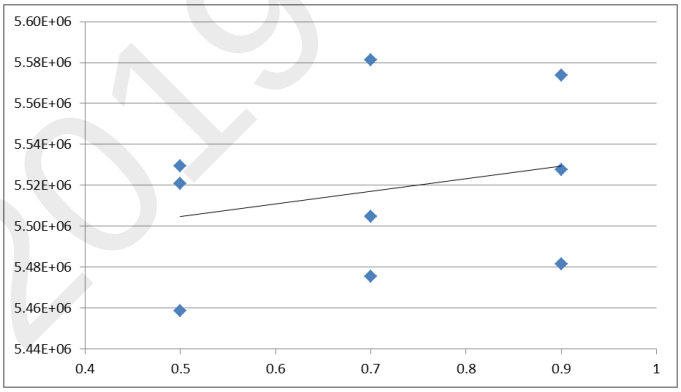


Fig. 9. Crossover probability value effect on average hyper-volume covered by the obtained nondominated front.

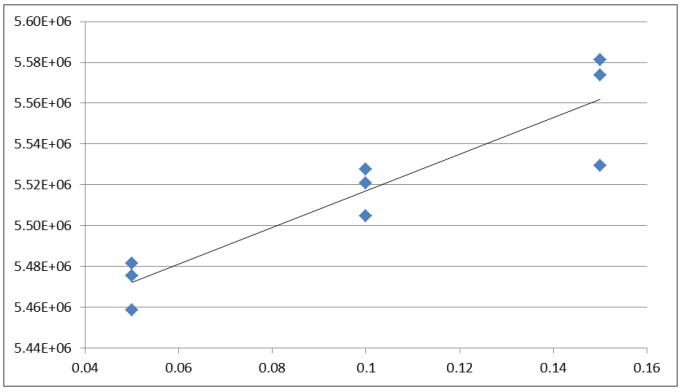


Fig. 10. Mutation probability value effect on average hyper-volume covered by the obtained nondominated front.

Fig. 11 and 12 provides the effect of the crossover and mutation probability values on computational time (in seconds) taken by the algorithm for each test problem. One can observe a positive correlation between the probability values of

crossover and mutation and the time required by the algorithm to converge. By tuning these probabilities (especially the crossover probability) to low values the exploratory propensity of the search declines and the algorithm is more likely to quickly be trapped in a local optimum. On the other hand, using high values discourages exploitation of good regions of the solution space and makes the algorithm take more time to converge. The crossover probability seem to have a more significant impact on the average computation time which is understandable since having more offspring in a population rather than replicating the old individuals is more expensive from an efficiency point of view.

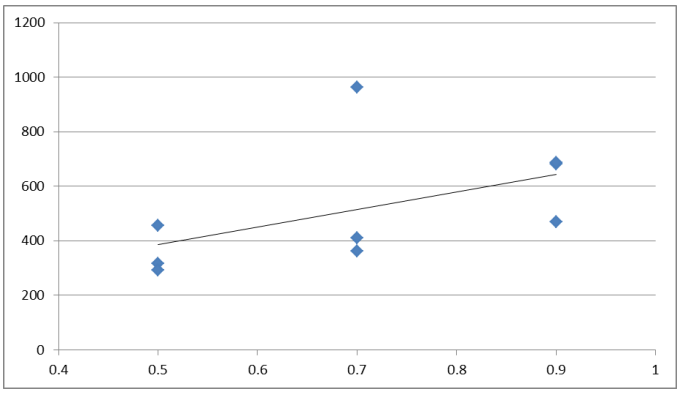


Fig. 11. Crossover probability value effect on average computation time.

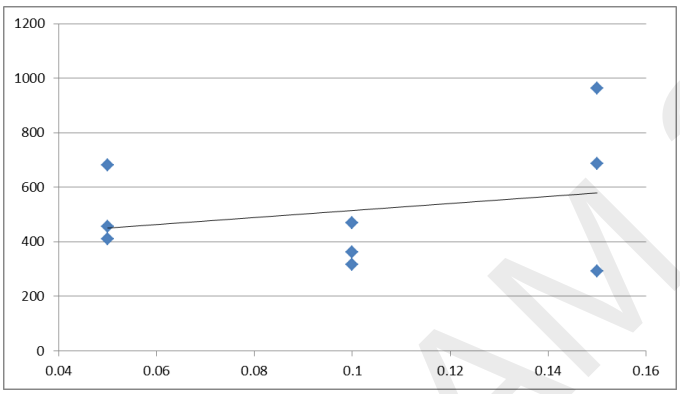


Fig. 12. Mutation probability value effect on average computation time.

V. Conclusion

This paper is new step forward into a new area of study that has not been yet addressed properly in the literature of sustainable production scheduling. The results given by the Non dominated Sorting Genetic Algorithm II suggests the possibility of using multi-objective metaheuristic algorithms and artificial intelligence techniques in order to better approximate the Pareto optimal frontier between conflicting objectives such as the ones considered in this paper. We look forward to enrich this ongoing study by enhancing the performance of the NSGA-II through considering schedules where planed delays of operations are permitted or when the nonresumability assumption is omitted. Also the results of this research suggests that having an elaborate coordination between production and maintenance services may bring more benefits by further reducing the power consumption.

References

[1] EIA. In: International Energy Outlook 2018.

[2] EIA. Annual Energy Review 2009. Report No. DOE/EIA-0384(2009) [Release Date: August 19, 2010].

[3] EIA. Energy consumption by manufacturers – data tables. [accessed 12/05/2009].

[4] Fang, K., Uhan, N., Zhao, F., Sutherland, J., 2013. Flow shop scheduling with peak power consumption constraints. Ann. Operations Res. 206 (1), 115–145.

[5] Babu, C.A., Ashok, S., 2008. Peak load management in electrolytic process industries. IEEE Trans. Power Syst. 23 (2), 399–405.

[6] Bruzzone, A.A.G., Anghinolfi, D., Paolucci, M., Tonelli, F., 2012. Energy-aware scheduling for improving manufacturing process sustainability: a mathematical model for flexible flow shops. CIRP Ann. Manuf. Technol. 61 (1), 459–462.

[7] Fang, K., Uhan, N., Zhao, F., Sutherland, J.W., 2011. A new approach to scheduling in manufacturing for power consumption and carbon foot print reduction. Journal of Manufacturing Systems., http://dx.doi.org/10.1016/j.jmsy.2011.08.004.

[8] Talbi, E., (2009) Metaheuristics: from design to implementation. John Wiley & Sons, Hoboken.

[9] Liu, C., Yang, J., Lian, J., Li, W., Evans, S., Yin, Y., 2014. Sustainable performance oriented operational decision-making of single machine systems with deterministic product arrival time. J. Clean. Prod. 85, 318e330.

[10] Drake R., Yildirim MB., Twomey J., Whitman L., Ahmad J., Lodhia P., 2006. Data collection framework on energy consumption in manufacturing. In: IIE Annual Conference and Expo.

New Solution for Cold Start Problems in Recommendation Systems

1st Saida Souilah
Computer Science departement
Université 8 Mai 1945 Guelma
Guelma, Algeria
souilahsaida1995@gmail.com

2nd Mohammed Chaoui
Computer Science departement
Université 8 Mai 1945 Guelma
Guelma, Algeria
chaoui.mohammed@univ-guelma.dz

***Abstract*—Contextual recommendation systems attempt to address the challenge of identifying the best-matched themes that best meet the needs of users by adapting to current contextual information. Many of these systems have developed in areas such as movies, games and music ... etc. One of the major problems with recommendation systems is to provide recommendations a new element to a new user or after finding a target user. This problem called cold start method; it is due to a lack of information about this entity and is a very important problem to deal with. In personal products like movies and music ..., passion of the users plays a surprising role in making decision, because the traditional user program model does not account for the impact of the user's passion. In addition, vendor of recommendation systems cannot understand and record users' ever-changing preferences.**

In this article, we introduce a system of recommendations of movies based on emotions to solve this problem. The goal of our system is to provide users with personalized and personal suggestions based on their preferences, as well as their feelings from the facial expression analysis that we use as a contextual parameter in conjunction with collaborative filtering and content-based filtering. For this, we expect users to be more satisfied with the recommendations done.

Keywords—cold start method, recommendation systems, emotion, preferences, and users.

I. INTRODUCTION

Network information is very important because often users cannot distinguish related information to unrelated information. This versatility has led to the use of referral systems to make it easier for users, who have become an indispensable part of solving the problem of information overload by providing users with interesting information on the Internet. Most current referral systems rely primarily on historical classifications and revisions to allow users to formulate recommendations, which typically face data volatility and start-up problems, due to the large number of items.

Consideration of context has introduced in recommendation systems to ensure users' needs in the short and long term, taking into account not only the history of preferences, but also their current situations. Previous studies of recommendation systems have examined the integration of emotion as a contextual parameter in the recommendation process. Emotions are mental states usually caused by an event of importance to the subject and have been modeled in various studies. The universal model classifies emotions into definite categories such as Ekman's six basic emotions (happiness, anger, fear, sadness, disgust and surprise). Emotions play an important role in rational and intelligent behaviors. The attention paid to the feelings of others allows us to understand each other better.

In this work, we will try to solve the problem of the cold start method of the film recommendation system as an important element of the entertainment in addition to its ability to improve the mood of the viewer. According to the analysis of user's emotions and its use as contextual parameter, both with collaborative filtering and content-based filtering, the contextual information (emotion) will used directly in the recommendation algorithm of our system (Contextual Pre-filtering approach).

II. RELATED WORKS

A recommendation system is a specific form of information filtering that aims to present a user with elements that are likely to interest him, based on his preferences and behavior. Therefore, we try to predict your appreciation of an element to suggest what you will be most able to appreciate, the recommendation systems (RS) can classified into three main categories: Collaborative Filtering (CF), Based on content (CB), and hybrid approach that combines the two previous methods. This three-step process begins with collecting user information. Then, we create a matrix to calculate associations. Finally, we are able to make a recommendation with a high level of confidence. This recommendation divided into two broad categories, one based on the users and the other on the items that make up the environment [1].

On the other hand, in Content-based filtering the system tries to recommend items that match the user's profile. In this approach, each item defined by a set of attributes and their values. Items with similar values for their attributes considered approaching (similar). When a user assigns a good score to an item, the items that are close to them considered potential recommendations [2].

The CB and FC approaches have both positive aspects and disadvantages. Sometimes it can be difficult to organize and categorize certain content [3]. Another problem that may arise is excessive specialization [3]. If the system is only a programmer and recommends items that are very similar to the user's previous preferences, the user will never receive recommendations for other types of items. On the other hand, there is a cold start problem [5]. In this case, each new element has no classification, so it will be impossible to recommend these items to users.

All these problems led different authors to experiment with the hybrid RS that merged CF approach and CB to avoid some limitations of the two previous systems.

More recently, several authors have tried to use the context in RS, as information is part of the context if it influences an interaction between two entities: "context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered to be relevant to the interaction between users and users, including the user and application themselves "[6].

In this section, we present some context-sensitive recommendation systems:

A. Micro-profiling

The micro-profiling approach aims to recommend unknown songs / artists to a user. The type of recommendation system is pre-filtering because the data is filtered according to the context (time: time of day, day, month or year), in this approach, time represents the time of day, it will be divided into several segments based on the songs the user listens to most during each moment (time segment) of the day [7].

B. Sourcetone Interactive Radio

In 2004, Malcolm Goodman and Dr. Jeff Berger founded a recommendation system that they named Sourcetone interactive radio (www.sourcetone.com). This system allows you to experience music in a new way while continuing to recommend music according to what the user wants to feel. Where the system asks the listener to determine his emotional state before making a recommendation.

C. Amazon

Another context-sensitive recommendation system is Amazon's Jeff Bezos, which created in July 1995 and opened in 2000 in France. This system requires users to login by name to create their profile and provides a "find a gift" button for each user to distinguish between user-specific preferences and preferences of the person to whom he will offer this gift [8].

In addition, some researchers also agree that colors have the power to influence our emotions and that color is a natural form to represent human emotions. Many researches highlights the relationship between colors and emotions [7].

III. Analysis of feelings

An emotion is a transient affective reaction of more or less intensity, which occurs in response to a triggering event. The American psychologist Paul Eck-mann, a pioneer in the study of emotions, observed facial expressions in various cultures and enumerated six fundamental emotions: joy, anger, fear, sadness, surprise, disgust.

They serve as a basic material for the development of other so-called secondary emotions. Emotional vocabulary describes the palette an emotion is an indicator of what is happening in us. Identifying it and taking into account the information is useful to act afterwards [9].

In the table below Table 1, we summarize the six emotions according to the psychologist Paul Eckmann:

TABLE I. Paul Eckmann Emotions

Emotion	Definition
Joy	Is linked to the satisfaction of a desire, It is a state of satisfaction and well-being that is manifested by gaiety and good humor. It increases our energy, motivation and self-confidence.
Anger	Is a protective reaction. It results from a frustration, a feeling of injustice, the meeting of an obstacle, even the attack on its physical or psychological integrity.
Fear	Is an anticipatory emotion. It is useful when it informs us of a danger, a potential or real threat because it prepares us to flee, or to act. It can also be related to apprehension, so it can be stimulating or blocking.
Sadness	Is linked to a loss, a disappointment, a feeling of helplessness, an unsatisfied wish. It is characterized by a drop in energy, motivation.
Disgust	Is a rejection, a physical or psychological aversion to an object (food ...) or a person, perceived as harmful.
Surprise	Is caused by an unexpected event, suddenly, in connection with an imminent change or by a revelation going against our perception, our representations. It is usually brief, then fades or gives way to another emotion.

Watching movies is not only fun, it can be a direct or indirect message that makes you think of many things, personal or public, to give you a beautiful image of the world even if it is slightly distorted, hope to face the future or at least help you realize your dream.

Can you adapt the films to our different feelings, positive or negative, and provide solutions?

The music, the lighting, the image and the dialogue of the film arouse a lot of emotion in the spectator, which leads to internal reflection, the display of problems and the resolution of problems. Some psychologists have sought to use these feelings to treat patients individually, collectively, or at the family level.

Films allow the viewer to explore his feelings and increase his communication skills, which many psychologists must identify as part of the treatment: the therapist therefore selects specific films that the patient can see alone or in person. Group, and then arrange a session with the therapist to discuss the feelings of the film and the similarities between this film and the patient's life.

The film arouses many emotions in the viewer: fear, surprise, enthusiasm, sadness and joy, offers solutions to problems, offers many therapeutic benefits, all in the entertainment environment and ensures the safety of some people who are afraid to talk directly about their feelings.

According to Gary Soloman, professor of psychology at the Community College of Southern Nevada, cinematography is effective in all psychological conditions and is an effective way for an individual to help himself. Solon Alex Heig, a member of the Royal College of Psychiatry, is a way for people who cannot express their feelings easily. It was a useful tool for psychology counseling.

The user profile is probably the key element of any information filtering system. Its acquisition, modeling, representation and evolution over time are major factors in the success of such systems [10], [11].

IV. Our approach

In our recommendation system, In addition to the user's preferences, we take into account the context (emotion) with which the user interacts with the system, which will improve the relevance of the recommendations.

The structure of our system contains six modules:

A. User profiles

The unit contains information about the personal data (name, password, age, sex) and the emotions of the user (emotional state of the user). To connect emotions to movies, we adapt the content of each film to the age and emotions generated for each user. In addition, the system captures the user's emotions by asking them to take a personal photo with a webcam while recording while ensuring the security of the information.

B. List of films

Contains information about recommended films and their characteristics (title, genre, director, actors and actress, year, rating).

C. The database

Used to create and update the user profile. It also interacts with the recommendation module to list recommendations.

D. The recommendation module

Searches for similar users and then recommend the most appropriate movies for the target user to his or her emotion. This module uses two recommendation techniques: collaborative filtering and content-based filtering.

E. The Emotion Detection module

Detects the emotions of the current user according to facial expression captured by the web cam.

F. User interface

For the interaction between the user and our system (registration, see recommendations, rating …).

V. RECOMMENDATION ALGORITHM

We divided our proposal into two parts:

A. Case 1: First start (1st phase)

Lack of information about the user's expectations through referral systems, particularly when recorded, can reduce the quality of referrals and their suitability for user preferences. This will quickly discourage this new user and make him give up the system.

In order to solve this problem, and since emotion is an important part of our system, when a new user is registered, we observe facial expressions captured by a webcam to extract their emotional state.

The derived emotion plays a role in the recommendations proposed by the system, where we rely on them to understand the case of the user and make a proposal that suits his psychological state when he uses the system.

1. Filtering based on emotion (1st iteration)

The idea, then, is to perform emotion-based filtering by integrating an attribute that presents it in the movies so that we can adopt the solution of psychologists who treat patients through movies in our system.

The result represented by a class that represents the emotional state of the user in question.

2. Age-Based Filtering (2nd iteration)

Age is a very important point in the categorization of films for users. For this, this step is mandatory in our system to eliminate inappropriate movies especially for children.

The result of this step is a new subclass of the previous class filtered according to the age of the user.

3. Our final classification (3rd iteration)

In this part, we propose to classify the last class based on age into subclasses as follows:

We divide the class into two subclasses:

- Sub class based on the number of views, the number of likes, the score, the year ... etc.
- Random class of the parent class.

We assign a percentage of 40% for the second class to answer the problem of recommendation of the items seen or never seen. Of course, respecting the interest of the user (another problem for new users) according to his emotional state that ensured in the first iteration.

In this first start, we answer the third need, which is the recommendation in the new systems.

The similarity integrated into the second phase, so that the recommendation personalized just for the user in question without offering him user-rated movies that are similar to him.

B. Case 2: Second start and up (2nd phase)

The idea is to integrate the two collaborative and content-based filtering in this phase with a percentage of 50% of a new class if there is a new emotion of the user in question (the 50% is the result applying the first iteration).

Otherwise, this class represents the 100% of the final recommendation starting with the resulting class by integrating age-based filtering only without the emotion.

VI. METHODS IMPLICATED

A. Collaborative filtering

The similarity algorithm applied to a set of profiles and returns a list of users deemed most similar to a given user. The recommendations based only on the user profile, without considering the description of the content, and it assumed that users with the same behavior have the same interests.

Will contain all the evaluations given by the user, they are collected as the user evaluates films.

The evaluation will be represented by a vector of Triplets (Film, Evaluation and Date) and the set of evaluation vectors of all the users constitute the matrix Users * Film.

This evaluation given directly by the user in the form of a score on a scale of 0 to 5, where:

1. Very weak;
2. Low;
3. Medium;
4. Good;
5. Very good.

In our work, we choose the Pearson correlation for the similarity calculation.

The similarity calculated by the following formula:

$$Sim(i,j) = \frac{\sum_{u \in I}(R_{a,i} - \bar{R}_a)(R_{u,i} - \bar{R}_u)}{\sqrt{\sum_{u \in I}(R_{a,i} - \bar{R}_a)^2}\sqrt{\sum_{i \in I}(R_{u,i} - \bar{R}_u)^2}}$$

Where:

$Sim(i,j)$: Is the similarity between the active user a and the user u.

I: The set of items having been noted by both user a and u.

$R_{a,i}$, $R_{u,i}$: Is the estimate of the item *I* by the user a and the user u respectively.

$\bar{R}_a$, $\bar{R}_u$: Average respective user notes a and u.

B. Filtering based on item content

In content-based recommendation systems, user profiles constructed from content information that the user views and evaluates. By example, if a user watches a movie and enjoys it, the system records the genres of the appreciated movie, in the user profile. If these genres already exist in the profile, their weights automatically be changed, afterwards, a matching between the user profile and the descriptor of a content is then performed [12].

The similarity between two items calculated by the following formula:

$$Sim(i,j) = \frac{\sum_{u \in U}(R_{u,i} - \bar{R}_i)(R_{u,i} - \bar{R}_j)}{\sqrt{\sum_{u \in U_i}(R_{u,i} - \bar{R}_i)^2}\sqrt{\sum_{i \in U_j}(R_{u,j} - \bar{R}_j)^2}}$$

Where :

$Sim(i,j)$: Is the similarity between item i and j.

U_i: The set of users who wrote for both items i and j.

$R_{u,i}$, $R_{u,j}$: The evaluation of the user u on item i and item j respectively.

$\bar{R}_i$, $\bar{R}_j$: The respective average of the notes of items i and j.

Once the similarity among the items has been calculated, the next step is to predict for the target user u, a value for the active itm i. One common way is to capture how the user evaluates similar items. The expected value based on the weighted sum of the user estimates as well as the deviations of the average estimates and can calculates using the following formula:

$$p_{u,i}(item) = \bar{R}_i + \frac{\sum_{k=1}^{K} sim(i,k) \times (R_{u,k} - \bar{R}_k)}{\sum_{k=1}^{K}(|sim(i,k)|)}$$

K : Number of items present in the neighborhood of item i, having already been noted by the user u.

$R_{u,k}$: User's note u for item k.

$\bar{R}_k$, $\bar{R}_i$: Average respective marks for items i and k.

C. Global Hybrid System Architecture (CF + CBF)

In addition to drawing on contextual information by combining content-based approaches and collaborative filtering when submitting the final recommendation.

The combination of these two algorithms improves the precision of the prediction,

It also gives more power to the scarcity recommendation system.

VII. Implementation and Results

Our method based on a method that allows giving results, according to user emotion that agrees to have at least first results that are appropriate to the user and eliminates the problems mentioned in the previous problem.

Our system is under implementation and as soon as we find results and user satisfaction, we will present them on our future papers.

References

[1] Charif Alchiekh Haydar, Les systèmes de recommandation à base de confiance, PhD Thesis, presented in Lorraine university, September 3, 2014.

[2] Toby Daigle, DÉVÉLOPPEMENT, INTELLIGENCE ARTIFICIELLE Introduction aux systèmes de recommandation, viewed March 2019.

[3] Balabanovi_c, M., Shoham, Y.: Fab: Content-Based, Collaborative Recommendation. Commun. ACM 40(3), 66{72 (1997).

[4] Billsus, D., Pazzani, M.J.: Learning Collaborative Information Filters. In: 15th Int. Conf. on Machine Learning. pp. 46{54. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1998).

[5] Schein, A., Popescul, A., Ungar, L., Pennock, D.: Methods and Metrics for Cold-Start Recommendations. In: 25th Int. ACM SIGIR Conf. on Research and Development in Information Retrieval. pp. 253{260. ACM, NY, USA (2002).

[6] Dey, Providing Architectural Support for Building Context-aware Applications.PhD thesis, Atlanta, GA, USA, A. (2000).

[7] E-MRS: Emotion-based Movie Recommender System Ai Thanh Ho Ilusca L. L. Menezes Yousra Tagmouti Department of Informatics and Operations Research University of Montreal, Quebec, Canada.

[8] Soltani Réda, Les systèmes de recommandation à base de confiance, Magister thesis, in Oran university, 2011.

[9] E-marketing, https://www.e-marketing.fr, Livre en ligne : La méga boîte à outils du Manager leader, Chapitre VIII : ÉQUILIBRE PERSONNEL, Fiche 07 : Identifier les 06 émotions fondamentales. DUNOD, Publié le 01/12/2017, dernière consultation le 10/12/2018.

[10] Saspost, https://www.sasapost.com, Article : Cinéma Guérir l'âme .. Les films peuvent-ils traiter nos maladies? Publié le 10/09/2017, dernière consultation le 10/12/2018.

[11] Reel Power: Spiritual Growth Through Film by Marsha Sinetar (Goodreads Author).

[12] Sarwar, B., Karypis, G., Konstan, J., and Reidl, J., "Item-based collaborative filtering recommendation algorithms". In WWW '01, 285-295, 2001.

Multiple classifier combination for steel surface inspection

Rachid Zaghdoudi
Research Center in Industrial Technologies CRTI
Algiers, Algeria
LabSTIC Laboratory
University of 08 may 1945
Guelma, Algeria
rachid_zmil@yahoo.fr

Hamid Seridi
LabSTIC Laboratory
University of 08 may 1945
Guelma, Algeria
seridihamid@yahoo.fr

Adel Boudiaf
Research Center in Industrial Technologies CRTI
Algiers, Algeria
a.boudiaf@crti.dz

***Abstract*—Vision-based steel surface inspection systems have gained increasing interest over the two past decades because they offer the possibility to meet the requirements of manufacturers in terms of time, cost and accuracy compared to traditional methods based on human vision. The main objective of this paper is to propose an efficient inspection system based on machine vision technology and multiple classifier combination to detect and classify the different types of defects in steel products. This system is based on two independent classifiers, support vector machine (SVM) and fuzzy K-nearest neighbor (FKNN). Features are extracted with the gray level co-occurrence matrix (GLCM) and Histogram of Oriented Gradients (HOG). Principal Component Analysis (PCA) is applied to these features in order to reduce the descriptor size and avoid over-fitting resulting from features redundancy. Each set of features is respectively inputted to SVM and FKNN to form four parallel classifiers. Also, seven fusion rules are applied to give the final decision. To evaluate the performance of the proposed system, a series of experiments was conducted on the NEU Surface Defects database. The results obtained demonstrate the effectiveness of the proposed approach for classifying steel surface defects.**

Keywords—Machine vision, classifier combination, surface defects, support vector machine (SVM), fuzzy K-nearest neighbor (FKNN), histogram of oriented gradients (HOG), gray level co-occurrence matrix (GLCM).

I. INTRODUCTION

Automated inspection of steel products by machine vision is one of the basic technologies for improving productivity and maintaining high product quality in many manufacturing industries. In fact, the integration of the industrial vision on a production line ensures a quality control of great precision, reliable, objective, fast and without contact. Any detected defect will cause the product to be marked or ejected, or even to stop the process. Several methods of steel products inspection under various conditions have been developed through numerous studies over the past decade, and the result of these studies has been used as one of the inspection tools applied to real systems. Inspection systems capture images of steel products with CCD cameras under specific lighting conditions; then various image processing algorithms are applied to the images and the defects of steel products are localized and classified using machine learning algorithms, these algorithms are crucial for the inspection systems. Fig. 1 illustrates the general structure of a product inspection system based on artificial vision technology.

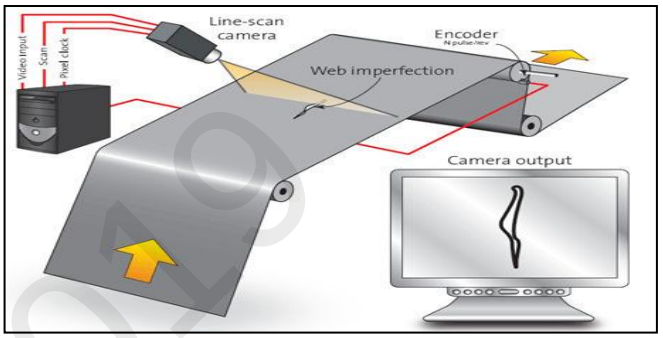


Fig.1. Surface inspection system of steel strip based on machine vision.

Different types of defects can be occurred during the steel strip production due to various causes. These defects do not only affect the visual appearance of the product, but they can also decrease its essential performances in varying degrees, such as corrosion resistance, abrasive resistance and fatigue strength. Therefore, it's very important for the inspection systems to classify these defects into various categories in order to adjust the production line and improve the quality of steel products.

In the last decade, many research studies have investigated the machine vision technology in the classification of surface defects for different types of steel products in different application environments. A number of surveys [1-2] on various aspects of automated steel products inspection have been reported. Hu [3], proposed a support vector machine model (SVM) to classify five defects in steel strip surface images using three types of features to represent the defect information, including geometric feature, gray scale feature and shape feature. Experiment results show that the proposed support vector machine model has the higher average classification accuracy than the one based on traditional back-propagation neural network .

Local Binary Pattern (LBP)-based methods and its various variants are widely used in the field of steel products inspection. Amid [4] describe a method to classify four common defect types in steel industry by using new local binary pattern (LBP) algorithm as feature descriptor, and employed decision tree and SVM with different kernels as classification method.

K. Song [5] proposed a new feature descriptor against noise named the adjacent evaluation completed local binary patterns (AECLBPs) to classify six famous different types of surface defects in hot-rolled steel strip products.

Q. Luo [6] built a novel system for steel surface defect classification based on generalized completed local binary patterns (GCLBP) framework. According to the authors the new descriptor differs from the variants of conventional local binary models, by its ability to innovatively excavate the implicit descriptive information from non-uniform patterns.

The authors of [7] developed a hybrid chromosome genetic algorithm (HCGA) combined with SVM kernel function selection, visual feature selection and SVM parameters optimization to establish the SVM multi-class model to classify five typical types of surface defect image collection obtained from the strip steel product line.

Another technique widely used in various pattern recognition applications is the Multi-scale geometric analysis (MGA) method. A system for surface defect classification of hot-rolled steels based on the Multi-scale geometric analysis (MGA) implemented via Curvelet transform and kernel locality preserving projections (KLPP) was proposed in [8]. According to the authors, this method has the characteristics of multi-scale and multi-directional analysis and is able to carve target entities on several multi-scale directional sub-bands.

An effective hybrid feature extraction method, called DST-KLPP [9], is proposed to take the advantages of both Shearlet transform and KLPP (Kernel Locality Preserving Projection). This method is applied to the classification of surfacedefects for 3 different types of metals. the experiments carried out showed that DST-KLPP can produce higher classification rates than the traditional methods such as Gray Level Co-occurrence Matrix and Wavelet transform, and also outperforms the other Multi-scale geometric analysis (MGA) methods such as Curvelet transform and Contourlet transform.

Ashour [10] developed a system of Surface Defects Classification of Hot-Rolled Steel Strips based on combining the use of discrete shearlet transform (DST) and the gray-level co-occurrence matrix (GLCM). The feature vector is formed by extracting multidirectional shear characteristics from each defect image. Followed by GLCM calculations from all extracted subbands, from which a set of statistical characteristics is extracted. Then the principal component analysis (PCA) is performed on the resulting feature vector in order to reduce the descriptor size and avoid over-fitting resulting from features redundancy. A multi-class support vector machine classifier is finally formed to classify six surface defects.

In the last few years, deep learning emerged as one of the best techniques in terms of accuracy for a large number of image processing tasks such as image segmentation, object detection and classification. In [11], A steel defect classification system based on Max-Pooling Convolutional Neural Networks (MPCNN) is proposed, which is able to learn, in a supervised manner, the visual features that achieve the best classification accuracy directly from the pixel representation of the steel defect images. W. Chen[12] proposed an ensemble approach to classify six different defects of hot-rolled steel strip surface. The proposed system consists of three different deep convolutional neural networks (DCNN) models trained individually and an average strategy is used to combine the output of them. The authors concluded that the proposed approach has made a state-of-art performance for steel surface inspection problem in terms of accuracy. But they mentioned that this method suffers from some drawbacks such as: computation cost and it doesn't deal with the case of presence of noise and low-quality of defect image.

In the field of pattern recognition, a large number of classifiers and feature extraction methods are available today. However, and despite the advances in this field, we realize that there is no classifier who can manifest an incontestable superiority over the other classifiers in all problems and all situations. Rather than seeking to optimize a single classifier, by choosing the best features for a given problem, the researchers found it more interesting to combine recognition methods. In fact, inspired by our work for off-line handwritten Arabic word recognition [13], the main objective of this paper is to propose an efficient inspection system based on classifier combination to classify six kinds of typical surface defects of the hot-rolled steel strip. This system consists of several phases: acquisition, preprocessing, feature extraction, classification and combination. After the preprocessing phase, features are extracted by the gray-level co-occurrence matrix (GLCM) and the histogram of oriented gradient (HOG). In addition, Principal Component Analysis (PCA) is applied to these features in order to reduce the dimensionality of the feature space. Then, each set is respectively inputted to SVM and FKNN to form four parallel classifiers. Finally, in the last stage, a set of rules is used to give the final decision.

The rest of this paper is organized as follows: in section 2, we will describe in more detail the architecture of the proposed inspection system. While experimental results are presented and discussed in section 3. The last section concludes the paper and outlines directions for future research.

II. PROPOSED SYSTEM

The main steps that composed the structure of the proposed system are illustrated in Fig.2 while each of these steps is described in detail in the following sections.

A. Preprocessing

The defect image is acquired by a camera or any other suitable digital device, and then it is submitted to the preprocessing phase in order to reduce the noise introduced by the digital device, remove useless information, and correct irregularities / variation. The preprocessing method used in the proposed system is the contrast enhancement, which is applied on the defect image in order to improve the quality. Histogram

equalization is a common technique for adjusting image intensities to enhance contrast. Fig.3 shows an example of the application of the proposed method for contrast enhancement.

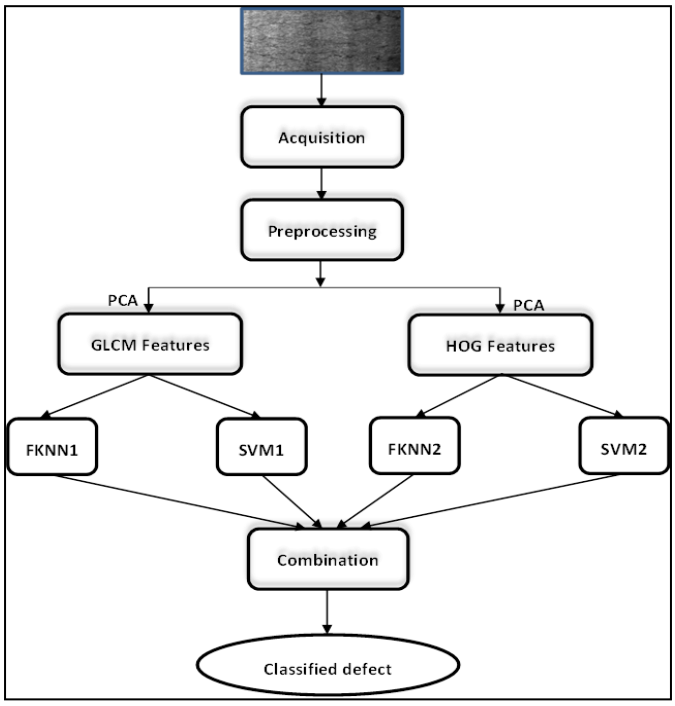


Fig.2. Overview of the proposed inspection system.

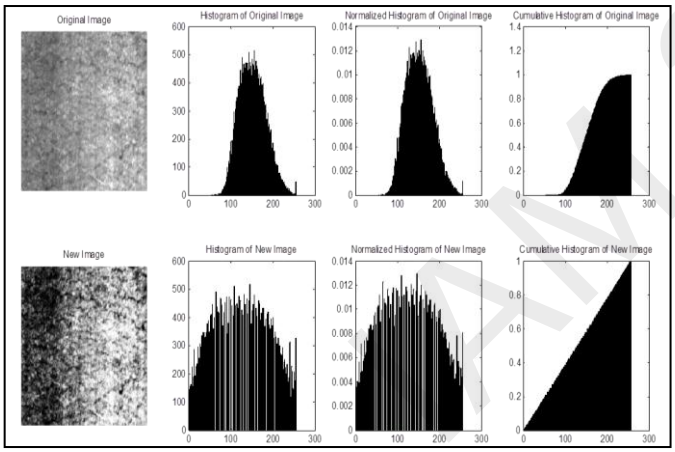


Fig.3. Contrast enhancement of defect image.

B. Feature extraction

After applying the preprocessing module to the input image, we must perform a feature extraction step to represent the surface defect image as a fixed-size feature vector. For our problem of steel surface defect classification, we have chosen to represent the defect images by discriminative features. For each defect image, we extract a set of features based on the Grayscale Co-Occurrence Matrix (GLCM) and the Histogram of Gradients (HOG). Each of these features is discussed in the following sections.

1) Gray Level Co-occurrence Matrix (GLCM) method

This method also called as Gray level dependency Matrix, is considered by many researchers as the reference in texture analysis and is often used as a comparative method for new approaches. It is, in fact, simple to implement and offers good performance because of their richness in texture information. The method of co-occurrence matrices consists of studying the joint behavior of pairs of pixels spatially separated by a given translation. The amplitude of the translation does not generally exceed a few pixels, in order to take into account only very local neighborhood information. Rather than using gray level co-occurrence matrix directly as feature descriptor, a number of second order statistical texture features can be extracted from this matrix such as: Autocorrelation, Correlation, Energy, Entropy, Contrast, .Etc. In this study, 19 GLCM features are used to describe the steel surface defect image (see Table 1).

TABLE I. THE EXTRACTED GLCM FEATURES.

N°	GLCM features
01	Autocorrelation
02	Cluster prominence
03	Cluster Shade
04	Contrast
05	Correlation
06	Difference entropy
07	Difference variance
08	Dissimilarity
09	Energy
10	Entropy
11	Homogeneity
12	Information measure of correlation1
13	Information measure of correlation2
14	Inverse difference
15	Maximum probability
16	Sum average
17	Sum entropy
18	Sum of squares variance
19	Sum variance

2) Histogram of oriented gradients (HOG)

Histogram of Oriented Gradients (HOG) is a feature extraction method based on the gradient that was first proposed by Dalal&Triggs in the pedestrian detection framework [14]. Since then, it has been widely used in a wide range of computer vision applications such as: face recognition, hand gesture recognition, Arabic handwritten recognition. The histogram of oriented gradients descriptor (HOG) consists of describing the local object appearance and shape within an image by the distribution of intensity gradients or edge directions. The main steps involved in the implementation of the HOG descriptor algorithm can be summarized as follows:

Step1: Compute the horizontal and vertical gradient of the image using centered 1-D derivatives.

Step2: Divide the image into N*N cells, and for each cell computes a histogram of gradient directions or edge orientations for the pixels within the cell.

Step 3: Each cell is discretized into angular bins according to the gradient orientation.

Step 4: Each cell's pixel votes for an orientation between 0 and 180 in the unsigned case, or between 0 and 360 in the signed case. Votes were weighted by gradient magnitude.

Step 5: Formation of blocks by grouping adjacent cells in spatial regions.

Step 6: Normalize each block histogram.

Step 7: Form the final descriptor by concatenating the blocks histograms.

In addition principal component analysis (PCA) is applied to the HOG features to reduce the dimensionality of the feature vector, and we call this new feature PCA-HOG features.

C. Classifiers

Classification consists in determining the class of the defect image. Features extracted from the training examples are used to learn the classifier to distinguish between different classes. In our study, we aim to explore different parallel combinations of classifiers applied to steel surface defects classification with the objective of improving the accuracy rate. We have chosen a Support Vector Machine (SVM) and fuzzy k-nearest neighbors (F-KNN) as base classifiers. Features based on the GLCM and HOG are calculated from the images of the training base, will be presented to each of these classifiers, to form a total four base classifiers that are combined using different combination rules. The classification stage is well described and discussed in our previous work [13].

D. Classifier combination

The different classifier combination rules used in our steel defects inspection system include majority voting, rules: minimum, maximum, sum, average, product, and the Bayesian method. Each of these combination rules is briefly discussed below.

1) Majority voting: assigns the class on which the majority classifiers agree to the input pattern.

2) Minimum rule: looks for the minimum score among the outputs of each classifier and assigns the class with the highest score to the input pattern.

3) Maximum rule: looks for the maximum score among the outputs of each classifier and assigns the class with the highest score to the input pattern.

4) Sum rule: sums up the score provided by each classifier and assigns the label with the highest score to the input pattern.

5) Average rule: takes the average of the scores of each class provided by the ensemble of classifiers and assigns the highest score to the input pattern.

6) Product rule: multiplies the score provided by each classifier and assigns the class label with the highest score to the input pattern.

7) Bayes method: assumes that the classifiers are mutually independent. It is a statistical fusion method that can be used for combining the outputs at the abstract level by using the confusion matrices of the member classifiers.

III. EXPERIMENTAL RESULTS

The performance of the proposed steel inspection system is evaluated by conducting a series of experiments on the Northeastern University (NEU) surface defect database [5] comprising 1800 gray scale images: 300 samples each of six different classes of typical surface defects such as: crazing (Cr), patches (Pa), pitted surface (PS), inclusion (In), rolled-in scale (RS), and scratches (Sc).Sample images of six kinds of typical surface defects are illustrated in Fig. 4; the original resolution of each image is 200 * 200 pixels. For the experiments carried out, we used 1200 images for the learning stage (200 images for each type of defect) and 600 images for the test stage (100 images for each type of defect).We present the experimental results on individual classifiers followed by the results on different combinations of these classifiers.

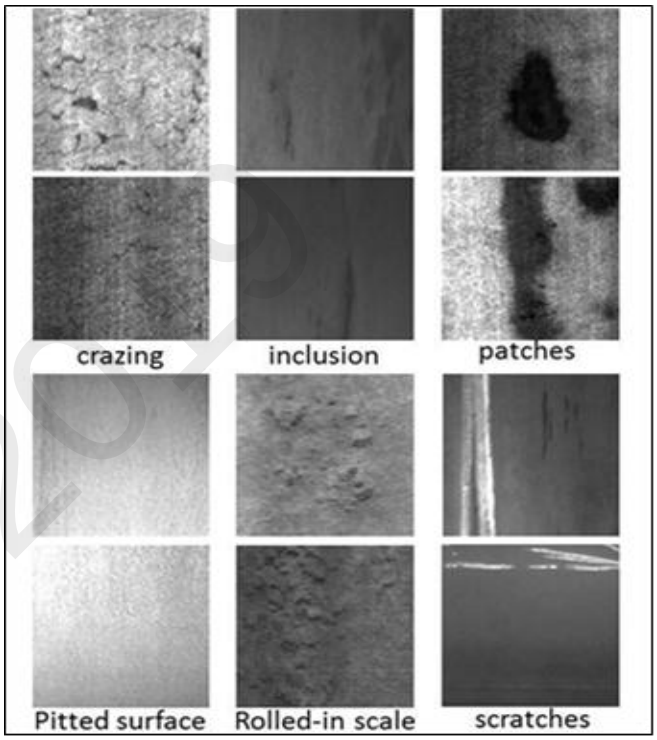


Fig. 4. Samples images of six classes of typical surface defects.

A. Performance of the individual Classifiers

The aim of the first experimental phase is to optimize the classifiers parameters and the feature extraction settings in order to get the best performance of the base classifiers, which will be then fused in the second phase of experimentation. As we have already indicated, features based on HOG and GLCM are separately fed to two different classifiers: SVM and F-KNN. The SVM classifier has two turning parameters: penalty (C) and RBF kernel parameter (γ) which are determined experimentally through a grid search using 5-fold cross-validation. While the FKNN classifier has a single parameter the number of nearest neighbors (k) which is also found experimentally. Table 2 summarizes the accuracy rates realized by the two classifiers for different number of PCA-GLCM features.

TABLE 2. ACCURACY RATES AS A FUNCTION OF THE NUMBER OF PCA-GLCM FEATURES.

PCA-GLCM	F-KNN (k=5)	SVM ($C=10^4$, $g=5*10^{-5}$)
19	83.33%	89.66%
15	83.16%	88.33%
12	82.67%	87.83%
10	82.33%	88.03%
08	**83.33%**	88.67
06	83.03	**89.66%**
05	81.13	89.03
04	80.37	86.17

The accuracy rates of the two classifiers for different number PCA-HOG features are summarized in Table 3.

TABLE 3. ACCURACY RATES AS A FUNCTION OF THE NUMBER OF PCA-HOG FEATURES.

PCA-HOG	F-KNN (k=6)	SVM ($C=5*10^3$, $g=8*10^{-4}$)
180	80.33 %	85.33%
162	79.67	84.83
144	79.16	84.67
126	80.16	84.83
108	79.67	84.83
90	79.83	85.00
72	**80.33 %**	85.16
54	79.00	**85.33**
36	66.67	82.33
18	56.16	67.33

Fig. 5 summarizes the highest accuracy rates achieved with the PCA-GLCM and PCA-HOG features using SVM and F-KNN classifiers. Among the two feature types, PCA-GLCM based features outperform the PCA-HOG features for each of the classifiers. Comparing the two classifiers, SVM achieve better accuracy rates on both the features. Overall, the highest accuracy rate of 89.66 % is achieved using PCA-GLCM features and SVM as a classifier.

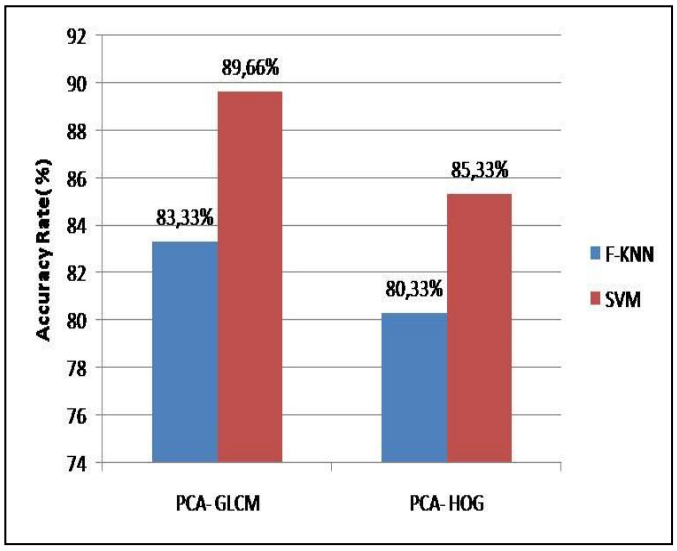


Fig.5. Highest accuracy rates of the two classifiers.

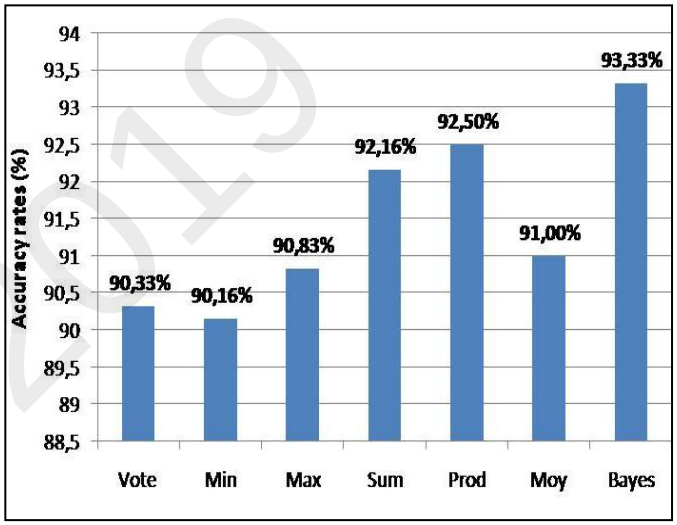


Fig.6. Accuracy rates on classifier combinations.

B. Performace of the Ensemble Classifiers

Once all classifiers have been established, the task now is to fuse theirs outputs under a suitable method of decision combination. For that, different parallel classifiers combination rules have been used and compared in term of accuracy rate. They include majority vote, minimum rule, maximum rule, sum rule, average rule, product rule, and Bayes method. Fig. 6 summarizes the performances of the ensemble classifiers obtained under different considerations.

It can be clearly observed from Fig. 6 that the classifier combination methods achieve better accuracy rates than the highest accuracy rate among the individual classifiers. The most significant improvement can be seen in case of Bayes method which realizes an accuracy rate of 93.33% as opposed to 89.66 % achieved with PCA-GLCM features and SVM classifier.

IV. CONCLUSION

In this study, we have presented a simple and efficient steel surface defects classification system based on multiple classifier combination. The system is composed mainly of two basic classifiers SVM and FKNN which use the same patterns to be classified from the (NEU) surface defect database. Four sets of features were extracted by GLCM and HOG from the training database. Each feature set was respectively inputted to SVM and FKNN to form a total four parallel classifiers. First, tests conducted on the four classifiers showed that the SVM classifier outperformed FKNN, regardless of the features that were used, either by GLCM or HOG features. The best accuracy rate which is equal to 89.66% was achieved when an SVM was paired with GLCM features. Second, the outputs of the four classifiers were fused using different combination methods such as majority vote, minimum rule, maximum rule, sum rule, Moy rule, product rule, and Bayes method. Tests carried out showed that Bayes method outperformed other combination schemes. The maximum score achieved is 93.33%, while the overall improvement varied from 0,50% to 03,67% compared to the greatest accuracy score of the individual classifiers. The experimental results showed that this method is simple and efficient for classifying the steel surface defects. Moreover, the proposed steel inspection system based multiple classifier combination provides a better results as well as more accuracy of 93.33%.

REFERENCES

[1] N. Neogi, D. K. Mohanta, and P.K. Dutta, " Review of vision-based steel surface inspection systems," EURASIP J Image Video Process. 2014(1), 1–19, 2014.

[2] X Xie, "A review of recent advances in surface defect detection using texture analysis techniques," Electron. Lett. Compu. Vision Image Anal. 7(3), 1–22, 2008.

[3] H. Hu, Y. Li , M. Liu, W.Liang, " Classification of defects in steel strip surface based on multiclass support vector machine," Multimedia Tools and Applications, v. 69 n.1,p. 199-216, March 2014.

[4] E. Amid, S. R. Aghdam, H. Amindavar, "Enhanced performance for support vector machines as multi-class classifiers in steel surface defect detection", World Academy of Science Engineering and Technology, vol. 6, no. 7, pp. 1096-1100, 2012.

[5] K. C. Song, Y. H. Yan, "A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects," Applied Surface Science, 285 (2013) B, 858‑864, 2013.

[6] Q. Luo, X. Fang, Y. Sun, L. Liu, J. Ai, C. Yang, O. Simpson, " Surface Defect Classification for Hot-Rolled Steel Strips by Selectively Dominant Local Binary Patterns", IEEE Access, vol. 7, pp. 23488 – 23499, 2018.

[7] H. Hu , Y. Liu , M. Liu , L. Nie, "Surface defect classification in large-scale strip steel image collection via hybrid chromosome genetic algorithm," Neurocomputing, v.181 n.C, p.86-95, March 2016.

[8] K. Xu, Y. Ai, X. Wu, "Application of multi-scale feature extraction to surface defect classification of hot-rolled steels," Int. J. Miner. Metall. Mater. 20 (1) , p. 37–41, 2013.

[9] K. Xu, S. Liu, Y. Ai, "Application of Shearlet transform to classification of surface defects for metals," Image and Vision Computing, v.35, p. 23–30, 2015.

[10] M.W. Ashour, F. Khalid, A. Abdul Halin, L. N. Abdullah, S.H. Darwish "Surface Defects Classification of Hot-Rolled Steel Strips Using Multi-directional Shearlet Features," Arabian Journal for Science and Engineering. v.4 n.4, pp. pp 2925–2932,April 2019.

[11] J. Masci, U. Meier, D. Ciresan, J. Schmidhuber, G. Fricout, " Steel defect classification with max-pooling convolutional neural networks," The 2012 International Joint Conference on Neural Networks (IJCNN), 2012.

[12] W. Chen, Y. Gao, L. Gao, X, Li, "A New Ensemble Approach based on Deep Convolutional Neural Networks for Steel Surface Defect classification," Procedia CIRP, v. 72, pp. 1069-1072, 2018.

[13] R. Zaghdoudi, H. Seridi, "Combination of Multiple Classifiers for Off-Line Handwritten Arabic Word Recognition," The international Arab Journal of Information Technology, Volume 14, No. 5, September 2017.

[14] N. Dalal, B. Triggs, "Histograms of Oriented Gradients for Human Detection," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893, 2005.